AD_____

AWARD NUMBER DAMD17-97-1-7306

TITLE: The Role of Endogenous Retroviral Sequences in Breast Cancer

PRINCIPAL INVESTIGATOR: John T. Welsh, Ph.D.

CONTRACTING ORGANIZATION: Sidney Kimmel Cancer Center
San Diego, California 92121

REPORT DATE: October 1998

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

19991020 063

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE October 1998	3. REPORT TYPE AND DATES COVERED Annual (22 Sep 97 - 21 Sep 98)

4. TITLE AND SUBTITLE
The Role of Endogenous Retroviral Sequences in Breast Cancer

5. FUNDING NUMBERS
DAMD17-97-1-7306

6. AUTHOR(S)
John T. Welsh, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Sidney Kimmel Cancer Center
San Diego, California 92121

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We have developed an approach to the identification of retroposed retroviral sequences in the human genome. This method uses nested PCR directed toward the LTRs of endogenous retroviruses, and arbitrary priming to complete the other end of the PCR product. Our results indicate that we can reliably identify sequences flanking retrovirus proviral sequences. This will allow us to probe BAC and YAC arrays with probes derived from tumors to determine whether the tumor displays the movement of a retrovirus between chromosomal positions. This array-based method will also lend itself to more facile examination of the products that result from RDA and SSH subtraction methods. Thus, we are now armed with an arsenal of tools for tracking retroviral retroposition events associated with human malignancy.

14. SUBJECT TERMS Breast Cancer			15. NUMBER OF PAGES 23
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

_____ Where copyrighted material is quoted, permission has been obtained to use such material.

_____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

_____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

_____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

_√_ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

_√_ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

_√_ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

_√_ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

_____ _____
PI - Signature | Date

12/8/98

Introduction

HERVs which make up an estimated 1% of the entire human genome are grouped into single and multiple copy families. Although most of them seem to be defective due to multiple termination codons and deletions, some of them are full-length proviruses and transcriptionally active. HERV transcription may play a role during normal gene expression (some transcripts are very abundant in placenta) and it has been suggested that this role could be a protective function against superinfection by related exogenous retroviruses.

However, overexpression of HERVs has been reported in various cancer cell lines, such as teratocarcinoma cell lines and cell lines derived from testicular and lung tumors. Expression of HERVs was shown to be associated with mouse strains susceptibility to lung cancer. In our laboratory, we have found a variation in the expression of different family members of HERVs in some colorectal cancer cell lines. The implications of expression of HERV sequences in pathophysiological processes remains to be elucidated.

There is substantial evidence that HERVs are still retrotransposing. As the retroviral long terminal repeat (LTR) sequences of HERVs contain complex regulatory elements such as promoters, enhancers, transcription initiation sites and polyadenylation signals, their insertion in the vicinity of host gene could cause a dramatic change in its expression. The current project relies on the hypothesis that HERVs have a role as mutagenic agents in at least a subset of breast cancers. This research focuses on two different families of HERVs: HERV-H family, that is the most abundant (1000 copies in addition to a similar number of solitary LTRs), and HERV-K family, that is characterized as the most "active" family. We are looking for differential HERV reverse transcriptase activity in cancer lines in response to growth factor and hormone treatments as well as in fresh breast cancer tissues. We are trying to determine if integration of cDNA occurs at an enhanced rate in cancer cells by looking for new integration sites of HERV sequences in tumor tissues. Thus, the purpose of this study is to develop ways to identify new insertional events that might result in breast cancer, and to gather other evidence that retroposition of endogenous retroviruses is an active mechanism of cancer.

Body

We have made considerable progress in this study, particularly in the refinement of the tools necessary for discovering retroposition events that, even at an enhanced rate, must be distinguished from the numerous proretroviral sequences in the genome. Here we discribe this progress, as well as a related study wherein we ask whether retrovirus reverse transcriptase is overexpressed in some cancers.

Elevation of Reverse Transcriptase Activity

Several instances of elevated levels of reverse transcriptase have appeared in our screens of normal vs. tumor pairs, and in cancer-derived cell lines. This is consistent with the idea that retroviral behavior is significantly altered in this subset of tumors, but we have not yet uncovered the mechanism.

RNA was isolated from two normal cell lines, eight different cancer cell lines, and from normal and cancer tissues originating from anonymous patients. Reverse transcriptase activity was measured through the abundance of retroviral reverse transcriptase mRNAs (HERV-H family) in the different cell lines and normal-tumor pairs.

Figure 1 shows the results of RT-PCR directed toward retroviral sequences for normal-tumor pairs, and for cancer derived cell types. Two tumors (lanes 19, 21) as well as two tumor derived cell lines (12, 15) show elevated levels of transcript for reverse transcriptase. This behavior is consistent with what would be expected for a retroposition mechanism. These samples are now being tested for elevated retroposition activity.

Figure 1. Retroviral Reverse Transcriptase transcript overexpression in cell lines and tumors, analyzed by RT-PCR.

1-2, cancer cell lines; 3-8, normal-tumor pairs; 9-10, normal cell lines, 11-16 cancer cell lines; 17-20, normal-tumor pairs; 21, unpaired tumor; 22-23, normal-tumor pair. Actin control and retroviral reverse transcriptase bands are indicated.
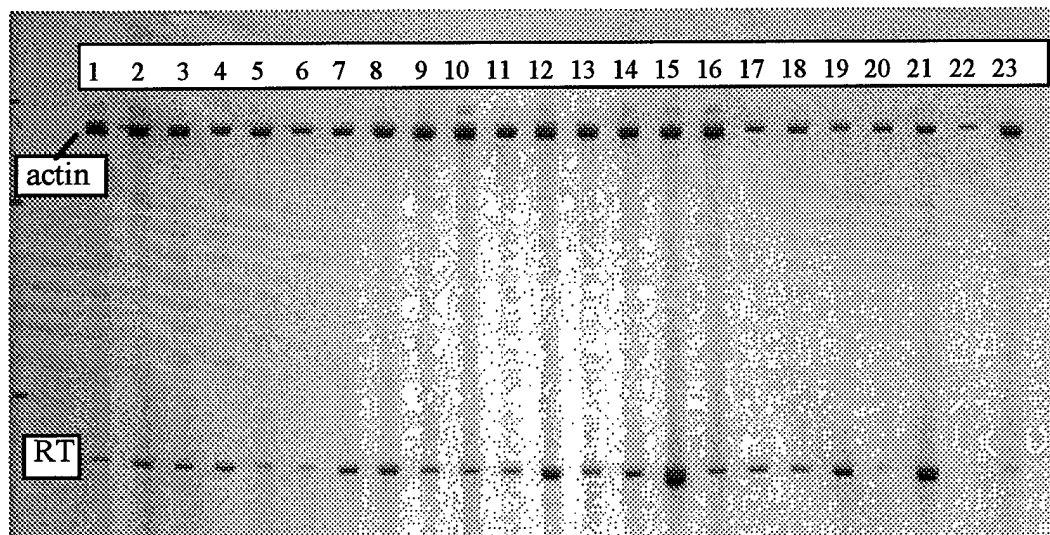
PCR of flanking genomic sequences.

PCR was performed using a primer specific for the LTRs of members of the HERV-H family (RTVL-H2, RGH-1, RGH-2). Specific targeting of HERV LTRs was confirmed by the use of nested primers generating products 10 basepairs shorter (**Figure 2**). In the examples to follow, the first PCR step involved a primer specific for the LTRs of the members of the HERV-K family (K10, HML6), under low stringency conditions. The second primer was specific for the LTRs of the member of the HERV-H family (RTVL-H2, RGH-1, RGH-2). The PCR was run under high stringency conditions to ensure specific sampling of HERV-H LTRs. The protocol was elaborated using cancer cell lines, then five matching pairs of normal and tumor tissue DNAs were screened for differences between normal and tumor tissues. Only one difference was detected in these fingerprints. To enrich for differences in low abundant products that are not detectable on DNA fingerprints, we applied a new approach recently developed in our lab. This method uses RNA or DNA fingerprints as complexity-reduced probes to hybridize DNA arrays. Differences detected between the membranes hybridized with cancer tissue DNA fingerprints and the ones hybridized with normal tissue DNA fingerprints are currently identified to check if they represent the hypothesized new integration sites of HERVs.

Figure 2 contains LTR to arbitrary products for 5 normal-tumor pairs, with the original primer pair in the left panel and the nested pair in the right panel. Each sample is represented in two lanes, where input RNA concentrations were titrated by a factor of two. The effect observed (arrows point several of the many examples) indicates that the initial primer pair targets predominantly retroviral sequences in the genome, because nesting the primers with respect to the original primers according to known, conserved 3' sequence causes a 10 base pair shift consistent with the 10 base pair nesting from one end. Sequence analysis of several of these bands confirmed that these bands usually derive from an LTR on one end, and arbitrary priming at the other. This experiment raises the interesting possibility that genomic clone (BAC or YAC) arrays, which are now commercially available, could be used to discover new retroviral insertion sites. The reason for this is that this simple procedure allows us to get a sample of sequences flanking the retroviral insertion sites. Figure 3 shows an additional example, and also shows a (1) polymorphism, (2) possible retroposition event.

6

Figure 2. Nested PCR shows that retroviral sequences are targeted.

Normal-tumor pairs. Left: A primer directed outward from the LTR was used in conjunction with arbitrary priming to amplify sequences flanking retroviral sequences in the genome. Right: nested primer secondary amplification of the right hand fingerprint. Lanes 1,3,5,7,9 = tumor; Lanes 2,4,6,8,10, normal.
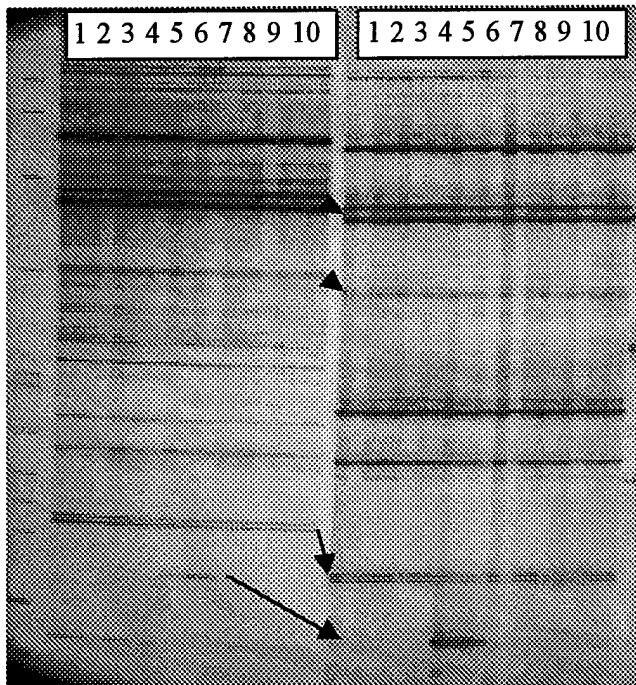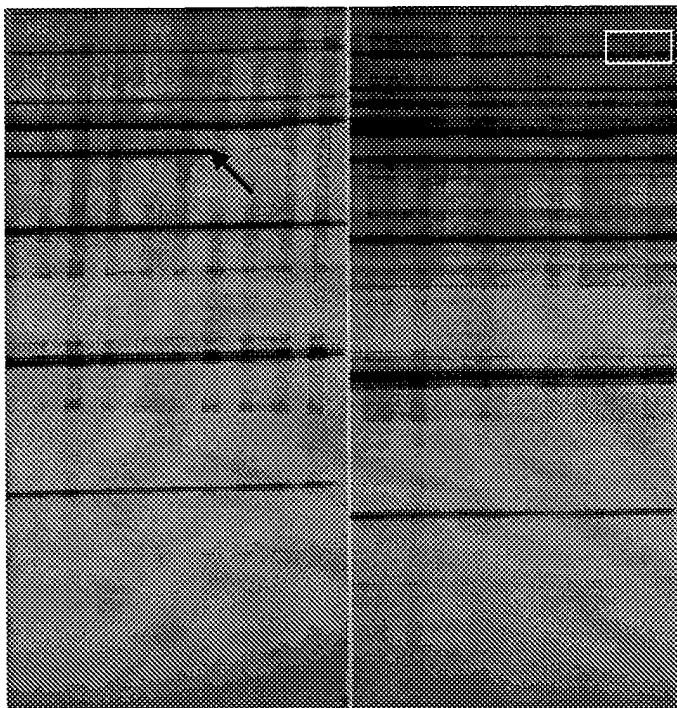

Figure 3. Additional example of LTR to Arbitrary sampling of retroviral flanking sequences.

This example also contains normal-tumor pairs, shows the shift due to nesting, and in addition shows both a polymorphism (indicated by the arrow) and a possible retroposition event (boxed).

We have succeeded in targeting PCR primers the HERV long terminal repeats and reverse transcriptase in various cancer cell lines and tumor tissues. A retroviral activity corresponding to different members of the two main HERV families was evidenced in cell lines and in tissues. No new integration site has been discovered yet but differences observed in normal versus tumor fingerprints and DNA array hybridizations are under further investigation (isolation and sequencing for identification). We are checking whether they represent the hypothesized new integration sites of HERVs and, if so, how they might contribute to tumorigenesis.

Our fingerprinting results so far indicate that the rate of retroposition is below 1 in two hundred. Using the array strategy and these same fingerprints, we should be able to push our sensitivity down to 1 in 4000, that is, we should be able to achieve 1 in 10 000 using only 10 primer pairs. This should allow us to identify novel retroposition events.

From our work on RNA fingerprinting, we know that in general fingerprints of the sort in Figures 2 and 3 contain many thousands of products that are below the detection capabilities of polyacrylamide electrophoresis. These products are present and reproducible, but not visible due to background and interference of other products. We discovered this approach to retroviral transposition while performing the necessary controls for the bubble PCR and RDA approaches originally proposed. When coupled with arrays, this approach may be more robust and easier to control than the bubble-PCR and RDA methods, and we plan to continue to explore this approach. There are two strategies the first being the amplification of subsets retroviral flanking sequences, using nested priming, and use these as probes for BAC and YAC arrays. In the alternative, we will try to use the Bubble-PCR and modified RDA approaches to enrich for retroposition events, followed by array analysis on BAC and YAC arrays.

Bubble-PCR, RDA, and SSH

Array methods may be necessary even for the bubble-PCR protocol originally proposed. There are clear advantages to the preliminary subtraction approach over the above fingerprinting approach, in that the fingerprinting approach requires iteration. However, the problem one encounters with bubble-PCR, RDA and SSH is that arbitrary priming is a very favored event when primers become as long as those necessary to facillitate bubble PCR. Thus, while the bubble PCR approach certainly enriches for retroviral sequences, one also amplifies many non-retroviral sequences. Normally, one can simply use higher stringency PCR conditions to achieve specificity, but this is not possible for the current project. This limitation derives from the fact that, while retroviruses can be highly conserved, they need not be, and the retroposition phenomenon we seek could occur with poorly conserved family members just as well as with well conserved family members. The same is true for our proposed modified RDA and SSH approachs. Thus, when fingerprints of bubble PCR, RDA or SSH amplified material are prepared, many of the products can come from non-retroviral sequences. While the gels in Figures 2 and 3 show a strong bias in favor of retroviral sequences, they really are very information poor, and any effort to increase the number of bands leads to the inevitable amplification of uninteresting sequences.

This is where arrays can help us. If we can mearly *enrich* for retroviral flanking sequences, arrays provide a way of purifying these sequences. We will spend some effort in making the method more robust. Our recent breakthrough in array-based cDNA analysis suggests itself as a potentially powerful tool for the study of this type of problem. In particular, we should be able to examine genes into which retroviruses have inserted by PCR amplification followed by cDNA array differential hybridization. This would serve the dual purpose of identifying an instance of a cancer-related retroposition, as well as identifying the gene that has been disrupted. Thus, we will try to amplify DNA flanking novel insertional events, and use available BAC and YAC arrays to identify regions of the genome where novel insertional events have occurred.

To facillitate this approach, we have conceived of a two-isotope phosphoimaging strategy for use with a phospho imager. The strategy is based on the idea that, because two isotopes of phosphorous decay at different rates, mixed probes with the control labeled with one isotope and the experimental labeled with another, can be hybridized to arrays simultaneously, thereby controlling for hybridization conditions and membrane array quality simultaneously. There is a great advantage in being able to use the exact same target spot for both hybridizations, because

variablity due to differences in spotting or colony growth disappears, making analysis much more reliable.

Other Parallel analysis methods

In addition to the above progress, we have continued to develop tools for understanding gene expression in general. We recently published on the use of RNA fingerprinting as a means of developing reduced complexity probes for array analysis. This method will allow us to detect alterations in gene expression due to transformation. The original outline of this grant proposed to identify retroposition events that might cause changes in gene regulation. We have also devised a method for analysing multiplex PCR amplified material on dot blots. Multiplex PCR is limited when gel electrophoresis is used as an analysis tool due to the complexities of the patterns that can result from all but the simplest multiplex reaction. The reason for this is clear, and the correct answer is not simply that resolution of gels is too poor. The real problem lies in arbitrary priming. Although we have had great success with various applications of arbitrarily primed PCR, in this instance it is a nusiance. Multiplex PCR reactions using more than about 10 primer pairs usually are seriously compromised. The complexity of primer sequences in the reaction support miss-priming at an exponentially greater rate than a pair of primers. The exponentiality is due to the fact that any miss-priming event from one primer can be paired with any upstream miss-priming event from any other primer. Arrays partially solve this problem, and may *completely* solve the problem. The array effectively purifies individual probe molecules from the reaction mixture, whether or not there are confusing, irrelevant molecules in the probe. Consequently, using this approach, a laboratory can probe arrays with probes directed toward any multiplex primer probe or twoard retroviral flanking sequences. A new retroposition event will be accompanied by new, unique flanking sequences which should show up as hybridization to a new BAC or YAC clone.

Conclusions

We now have a well-developed arsenal for the identification and characterization of retroviral sequences in the human genome, and have support tools that will be necessary to carry this project forward to the tumor screening stage. The advent of commercially available arrays promises to allow us to overcome some of the inherent limitations and experimental difficulties in standard bubble PCR, RDA, and SSH technologies. During the next few months we expect to be able to use these methods to screen tumors for rare retroposition events that may have causal implications in cancer.

References

McClelland, M., Mathieu-Daudé, F. and Welsh, J. (1995) RNA fingerprinting and differential display using arbitrarily primed PCR. T.I.G. 11, 242-246.

Mathieu-Daudé, F., Stevens, J., Welsh, J., Tibayrenc, M. and McClelland, M. (1995) Genetic diversity and population structure of *Trypanosoma brucei*: clonality versus sexuality. Mol. Biochem. Parasitol. 72, 89-101.

Mathieu-Daudé, F., Cheng, R., Welsh, J. and McClelland, M. (1996) Screening of differentially amplified cDNA products from RNA arbitrarily primed PCR fingerprints using single strand conformation polymorphism (SSCP) gels. Nucleic Acids Res. 24, 1504-1507.

Mathieu-Daudé, F., Welsh, J., Vogt, T. and McClelland, M. (1996) DNA rehybridization during PCR: the "Cot effect" and its consequences. Nucleic Acids Res. 24, 2080-2086.

Vogt, T., Mathieu-Daudé, F., Kullmann, F., Welsh, J. and McClelland, M. (1997) Fingerprinting of DNA and RNA using arbitrarily primed PCR. In: DNA Markers: Protocols, Applications and Overviews (Caetano-Anolles, G. and Gresshoff, P.M. Eds) John Wiley & Sons, New-York.

McClelland, M., Honeycutt, R., Mathieu-Daudé, F., Vogt, T. and Welsh, J. (1997) Fingerprinting by arbitrarily primed PCR. In: Methods in Molecular Biology, Differential Display Methods and Protocols (Liang, P. and Pardee, A.B. Eds) Humana Press Inc., Totowa, N.J.

Mathieu-Daudé, F., Welsh, J., Davis, C. and McClelland, M. Differentially expressed genes in the *Trypanosoma brucei* life cycle identified by RNA fingerprinting. Mol. Biochem. Parasitol. (in press).

Mathieu-Daudé, F., Trenkle, T., Jung, B., Vogt, T., Welsh, J. and McClelland, M. Identification of differentially expressed genes using RNA fingerprinting by arbitrarily primed PCR. In: Methods in Enzymology, cDNA preparation and display (Abelson, J.N. and Simon, M.I. Eds) Academic Press Inc., Orlando, FL. (in press).

Mathieu-Daudé, F., Benson, N., Kullmann, F., Honeycutt, R., Welsh, J. and McClelland, M. Screening of differentially expressed cDNAs using single strand conformation polymorphism (SSCP) gels. In: PCR Methods Manual (Innis, M., Gelfand, D. and Sninsky, J. Eds) Academic Press Inc., San Diego, CA. (in press).

Trenkle, T., Welsh J., Mathieu-Daude, F., Jung, B., and McClelland, M., Non-stoichiometric reduced complexity probesx for cDNA arrays. Nucleic Acids Res. 26: 3883-3891, 1998.

RNA Fingerprinting Displays UVB-specific Disruption of Transcriptional Control in Human Melanocytes[1]

Thomas M. M. Vogt, John Welsh,[2] Wilhelm Stolz, Frank Kullmann, Barbara Jung, Michael Landthaler, and Michael McClelland

Sidney Kimmel Cancer Center, San Diego, California 92121 [T. M. M. V., J. W., F. K., B. J., M. M.], and Department of Dermatology, University of Regensburg, Franz Joseph Strauss Allee 11, 93042 Regensburg, Germany [T. M. M. V., W. S., M. L.]

ABSTRACT

In mammalian cells, UV induces a limited set of early transcribed genes, which overlaps with the set of genes induced by tumor promoting drugs such as 12-*O*-tetradecanoyl phorbol-13-acetate (TPA). Among these are genes for transcription factors, the activation of which leads to complex secondary changes in expression of multiple target genes. How these delayed pleiotropic UV effects on transcription may contribute to initiation of melanoma skin cancer is poorly understood. We analyzed changes in the relative abundances of 1900 transcripts in newborn human melanocytes 8 h after treatment with UVB, TPA, and cycloheximide in all combinations, using RNA arbitrarily primed PCR for differential display. The relative abundances of 205 transcripts (11% of all transcripts surveyed) were altered by one or more of the treatment combinations. Fourteen of the 77 genes up-regulated by TPA were also up-regulated by UVB, but 60 of the TPA up-regulated genes were down-regulated by UVB, indicating both intersecting and independent signal transduction pathways for UVB and TPA. A number of UVB and TPA target genes were identified by cDNA cloning. Consistent with UVB induction of a partly transformed phenotype in mammalian cells, UVB antagonized the TPA-inducible expression of tumor-suppressive tropomyosin 3 mRNA. In addition, UVB may impair mitochondrial functioning and induce oxidative stress by strong down-regulation of mitochondrial transcription. Finally, increased expression of the dihydropteridine reductase gene, a major regulator of the cellular tetrahydrobiopterin pool, was linked to the UV pathway.

INTRODUCTION

The incidence of potentially fatal malignant melanoma skin cancer is increasing in many parts of the world, probably due to increased recreational UVB exposure (1). At present, malignant melanoma ranks seventh among all cancers in the United States with a total life-time risk of 1:87 (1). Details of the acute UVB effects in melanocytes are still poorly understood, because most investigations on UV have been done in *Escherichi coli*, yeast, and nonmelanocytic cells and have focused on the mutagenic and DNA-damaging action of UVC (2).

In mammalian cells, no straightforward signaling cascade analogous to the SOS system in *E. coli* has been identified (2, 3). The earliest detectable response to UVC in mammalian cells is the activation of SRC tyrosine kinase, followed by the activation of Ha-ras and Raf-1 near the cytoplasmic membrane (4–6). Stimulation of Ras leads to the activation of a cascade of cytoplasmic protein kinases (*e.g.*, JNK; Ref. 7), ultimately resulting in the activation of genes for transcription factors such as c-jun and ATF-2, components of AP-1 (8, 9), and *NF-κB* (10). Some growth factors, such as epidermal growth

factor, and tumor-promoting drugs, such as TPA,[3] lead to increased transcription of c-*fos*, also a major component of AP1, using a different set of protein kinases (*e.g.*, ERK) and thus intersect with the UV-regulated pathways, which also modulate c-*fos* and AP-1 activity (8, 11). DNA-activated nuclear protein kinases, which are inducible by UV damage, also alter the activity of transcription factors such as p53 and AP-1 (2, 12).

About 50 genes are known to participate in the mammalian UV response downstream of these early responses (13). Only a few of these are associated with mechanisms that may compensate for the genotoxic and oxidative effects of UV. Some of these genes are involved in cell cycle arrest in G_1 (14), a variety of DNA repair mechanisms (13), and *p53*-dependent programmed cell death (12). Many UV-inducible genes, however, reflect a partly transformed phenotype (2). Accordingly, UV-inducible genes such as c-*fos*, c-*myc*, *plasminogen activator, collagenase,* and *ornithine decarboxylase* are involved in tumor progression of a variety of cancers, including malignant melanomas (2, 13, 15, 16). Ornithine decarboxylase has recently been the focus as a target for new cancer therapies against malignant melanomas (17).

The available data indicate a complex and flexible network of intersecting signaling pathways in mammalian cells that integrate both stress, such as UV exposure, and signals from growth factors or tumor-promoting chemicals, such as TPA (2, 18). Few studies have addressed how UV-mediated disruption of transcriptional control in human melanocytes compares with the action of growth-promoting factors such as TPA at later time points. One way to characterize the UVB and TPA responses in melanocytes is to analyze melanocytic transcripts to discover those that are differentially regulated at later time points. Therefore, we analyzed abundance changes in a large array of transcripts 8 hours after treatment with UV, TPA, and CX in all combinations, using RAP-PCR (19, 20). RAP-PCR provides a complex molecular phenotype, the "RNA fingerprint," that allowed us to detect changes in the abundance of 205 of 1900 transcripts sampled after treatment with UVB, TPA, and CX in all combinations. The identification of late-transcribed target genes in melanocytes may provide new targets for tumor-preventive therapies as well as diagnostic and prognostic molecular markers for detecting UVB damage in precursor lesions associated with an increased risk of malignant transformation.

MATERIALS AND METHODS

Tissue Culture and Treatment Conditions. Cryopreserved human Caucasian newborn melanocytes (Clonetics Corp., San Diego, CA) were grown in a MCDB153 formulation supplemented with 1 ng/ml human fibroblast growth factor-basic, 0.5 μg/ml hydrocortisone, 10 ng/ml (16 nM) TPA, 5 μg/ml insulin, 50 μg/ml gentamicin, 50 ng/ml amphotericin-B, and 15 μg/ml bovine

---

[3] The abbreviations used are: TPA, 12-*O*-tetradecanoylphorbol-13-acetate; RAP-PCR, RNA arbitrarily primed PCR; RT-PCR, reverse transcription-PCR; CX, cycloheximide; SSCP, single-stranded conformational polymorphism; EST, expressed sequence tag; G3PDH, glyceraldehyde 3-phosphate dehydrogenase; OCR(s), observed category(ies) of regulation; TM3, human tropomyosin 3; DHPR, dihydropteridine reductase; nr, non-redundant.

pituitary extract. All chemicals and hormones were purchased from Clonetics Corp.

As a UV source, we used a customized apparatus built by Stratagene (San Diego, CA) with a built-in flux measurement device. The UV lamps provided by the manufacturer have a continuous spectrum of UVB light (290–320 nm) with a peak at 312 nm. The decrease of emission energy with shorter wavelengths is similar to sunlight; therefore, the emission in the UVC range (<290 nm) is negligible. About 15% of the total energy is emitted in the UVA range (320–400 nm). Mean UVB flux rates were 20 J/s. Overall, this lamp provides more physiologically relevant spectrum exposure than those generally used in comparable studies. Because the mammalian UV response has been studied most often using the 254-nm wavelength of UVC (200–290 nm; Ref. 2), an additional source emitting with a peak at 254 nm (germicidal UVC) was used for comparative experiments.

Human newborn melanocytes were grown in 100-mm dishes to 70–80% confluence over 8 days (doubling time of about 48 h). On day 7, the cells received growth factor-free media for 20 h. On day 8, the cells were treated with UVB, TPA, and CX in all possible (2^3) combinations: 0/0/0, 0/0/CX, 0/UV/0, TPA/0/0, TPA/UV/0, TPA/0/CX, 0/UV/CX, and UV/TPA/CX. All dishes were rinsed with sterile PBS; then the melanocytes in the UV treatment group were irradiated through a thin (1-mm) layer of PBS with a dose of 2000 J/m^2, as well as 1500 and 750 J/m^2. After irradiation, the PBS was replaced with fresh media supplemented or not supplemented with TPA (32 nM) and/or CX (0.02 mg/ml). All experiments were performed with two Clonetics cell lines HNME 680 and 2486. Cells were harvested 8 h after treatment for analysis of mRNA.

RNA Extraction. Total cellular RNA was extracted using the RNeasy spin column purification kit (Qiagen, Chatsworth, CA). To remove contaminating genomic DNA, the total RNA was treated with DNase I (0.2 unit/μl; Boehringer-Mannheim Corp., Indianapolis, IN) for 45 min at 37°C in a Tris-MgCl$_2$ buffer (each 0.01 M, pH 8.0) and in the presence of a RNase-inhibitor (1.2 units/μl; Boehringer-Mannheim Corp., Indianapolis, IN). After DNase treatment, another set of RNeasy spin columns was used for cleaning the treated RNA. The RNA concentration was measured spectrophotometrically at 260 nm and adjusted, and equal aliquots were then electrophoresed on 1% agarose gels stained with ethidium bromide to compare large and small rRNAs qualitatively and to exclude degradation. When starting with fresh RNAs, we performed one RAP-PCR, leaving out the reverse transcriptase as a control of DNA contamination.

RAP-PCR of Total Cellular RNA. RNA stock solutions at about 400 ng/μl were diluted to about 80, 40, and 20 ng/microliter. Five μl of the RNAs were combined with 5 μl of 2× first-strand cDNA reaction mixtures and incubated for 60 min at 37°C [final concentrations, 2 units/μl RNase inhibitor, 50 mM Tris-HCl (pH 8.3), 50 mM KCl, 4 mM MgCl$_2$, 10 mM DTT, 0.2 mM deoxynucleotide triphosphates, 2 μM first-strand arbitrary primer, and 18.75 units/reaction murine leukemia virus reverse transcriptase; Promega Corp., Madison, WI]. cDNAs were diluted 4-fold and subsequently cycled through 30 low stringency cycles (94°C for 1 min, 35°C for 1 min., and 72°C for 2 min) with 4 units/reaction Amplitaq DNA polymerase Stoffel fragment (Perkin-Elmer, Norwalk, CT), 4 mM MgCl$_2$, 0.2 mM each deoxynucleotide triphosphate, 2 μCi/reaction [α-^{32}P]dCTP, and 4 μM arbitrary second-strand primer.

Three μl of the complete reaction were mixed with 12 μl of formamide-dye buffer and denatured at 68°C for 15 min. Two μl of these solutions were loaded onto 8 M urea/6% polyacrylamide sequencing gels. Electrophoresis was performed for 4–6 h at 50 W in 1× Tris-borate-EDTA buffer. All three reactions at each starting concentration of RNA were loaded side by side. Gels were then transferred to 3 MM Whatmann paper, dried under vacuum at 80°C, and directly placed against X-ray film at room temperature. Several luminescent labels (autoradiogram markers) were stuck to the gel to facilitate alignment of the autoradiograms with the gels in case the isolation of one of the fragments was desired. Multiple exposures for a variable time (up to 24 h and more) followed, depending on the intensities of interesting bands.

Arbitrary PCR Primers. Primers were manufactured by Genosys (Woodlands, TX). For first-strand synthesis, the arbitrary primer AP-11 was used for all fingerprints (5'-AGGGGCACCA-3'). A variety of arbitrary second-strand primers (10–19 mers) was used to sample about 1900 transcripts from the experimental melanocytic RNA pools: AP-1, 5'-GAGGGT-GCCTT-3'; AP-2, 5'-GGTGCCTTTGG-3'; AP-3, 5'-CCAAATGCACCT-TCACC-3'; AP-4, 5'-GCACCAGGGG-3'; AP-5, 5'-GTGGTGACAG-3';

AP-6, 5'-AGGGGCACCA-3'; AP-7, 5'-AAGAAGAGCAA-3'; AP-8, 5'-ACGAAGAAGAG-3'; AP-9, 5'-CACCAGGGGC-3'; and AP-10, 5'-CCTA-CAAAGCTTATTCCTC-3'. All primer combinations were selected in previous experiments according to their ability to generate reproducible fingerprints with about 150 or more clearly visible products.

Isolation of Differentially Amplified PCR Products. For isolation of fragments that indicate differential gene expression, luminescent labels on the gel were aligned with their exposed images on the autoradiograms. The exact position of the fragment in the dried gel was marked with a needle. Gel slices (0.5–1 × 2–3 mm) carrying the target fragment were then excised with a razor blade and placed in 50 μl of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The DNA was eluted by incubating at 65°C for 3 h. The eluates were diluted with water 100-fold. Ten μl were taken for reamplification of the desired product using the primers and the conditions outlined above, for 20 PCR cycles, except that AmpliTaq DNA Polymerase (Perkin-Elmer) was used for this step. AmpliTaq DNA polymerase was used, because it is less tolerant of mismatch priming than Amplitaq DNA polymerase Stoffel fragment (Perkin-Elmer). The PCR products were routinely checked by denaturing PAGE running the reamplified product next to the original fingerprint to verify its size and purity.

Purification of the Desired Product by Native SSCP Gels. After a differentially amplified RAP product is detected, eluted, and reamplified, it is most often contaminated with a mixture of several different products of similar size. To identify the regulated transcripts, we used native polyacrylamide gels to separate the sequences of the reamplified mixture based on SSCP as described previously (21). The SSCP gels were run in 0.6× TBE for 14–18 h. We analyzed reamplifications of both the region carrying the desired product and a corresponding region from an adjacent lane, where the product was less prominent or not visible. Side-by-side comparison of these two reamplification mixtures on a SSCP gel allows the selection and reamplification of appropriate product for cloning, because contaminants can be identified and avoided (21). Once the correct band is identified by this procedure, it is cut from the gel and reamplified a second time. This results in isolation of a single product in more than 80% of the cases (21).

Cloning, Southern Transfer, and Sequencing of Interesting Products. Reamplified products from the SSCP gels were cloned using the Invitrogen TA cloning kit (Invitrogen, San Diego, CA). After blue-white screening of the clones, 10 white colonies per desired product were picked and suspended in 50 μl of water. Aliquots of these bacterial suspensions were checked by high-stringency PCR for the presence and the correct length of inserts using the universal and the M13 (−20) reverse sequencing primers. Clones with the desired inserts were subsequently grown overnight in 5 ml of LB medium containing 50 μg/ml of ampicillin for plasmid isolation.

Four to six clones/transcript were sequenced with the Applied Biosystems 373 automatic sequencer using the Perkin-Elmer (Norwalk, CT) DNA sequencing kit. The data bases of the National Center for Biological Information were searched to align the obtained sequences with known cDNA clones (GenBank nr), genomic clones (GenBank nr), and cloned expressed sequence tags (dbEST) using the two search modes blast/nr and tblastx/dbEST. Previously unknown sequences were submitted to GenBank as melanocytic ESTs.

If sequences were multiply represented and confirmed within the majority of the clones resulting from one RAP-PCR product, inserts were reamplified and used as probes against Southern blots of the original gels. This procedure insures that selected sequences were in fact differentially amplified in the original gels. Differential expression is then usually confirmed by Northern blot analysis or RT-PCR methods. The DNA from RAP-PCR fingerprints was transferred to nylon membranes (Duralon-UV; Stratagene) by capillary action overnight in a 20× SSC buffer. After UV cross-linking, the blots were prehybridized and hybridized with conventional methods (22).

Confirmation of Differential Expression of Selected Sequences by Northern Blot Analysis and Semiquantitative RT-PCR. Twenty μg of total RNA per treatment group was electrophoresed through 1.0% agarose/formaldehyde gels (4-morpholinepropanesulfonic acid buffer), transferred to nylon membranes (Duralon; Stratagene), and UV cross-linked. Probes were synthesized from the corresponding clones and randomly labeled with the prime-IT II kit (Stratagene). A *G3PDH* probe was used to verify that the lanes were equally loaded (Clontech Laboratories, Palo Alto, CA). Hybridization to blots was performed using conventional methods (22). Because Northern blots are sometimes not sensitive enough to confirm differential expression of transcripts of low abundance, we have recently developed a PCR-based method in
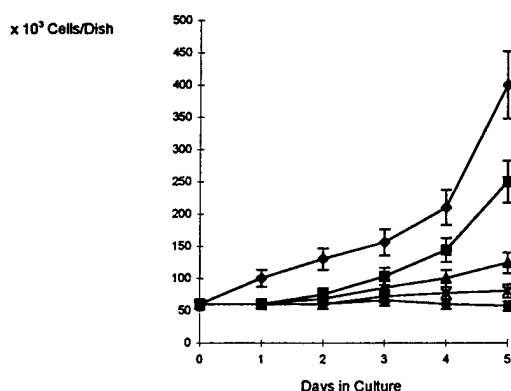
x 10³ Cells/Dish

Fig. 1. The growth of melanocytic cells is progressively suppressed by increasing doses of UV. A UV dose of 2000 J/m² (UVB, 312 nm) or 40 J/m² (UVC, 254 nm) results in complete cessation of growth for more than 3 days without dropping the viability to lower than 85%. ◆, no UV; ■, UVB 500 J/m²; ▲, UVB 1,000 J/m²; ×, UVB 2,000 J/m²; ●, UVC 40 J/m². *Bars*, SD.

our laboratory that allows estimation of the relative mRNA abundances.[4] This method uses a low-stringency RT-PCR protocol with specific primers, thus generating background fragments that serve as internal standards. To estimate relative mRNA levels of the dihydropteridine reductase gene transcript, the following primers were used: 5'-TGA CAG ACT CGT TCA CTG AGC AGG-3' and 5'-TCA TCG GGG TAT CCA GGG TAA C-3'.

Quantitation of the Northern and semiquantitative RT-PCR results was performed using the Personal Densitometer (Molecular Dynamics). The individual signals were corrected for uneven loading by division with the corresponding signal generated by the *G3PDH* probe or internal standard.

RESULTS

UVB-modulated Growth of Melanocytic Cells. Experiments were performed with UVB lamps, which better represent the natural spectrum of UV exposure of human skin than UVC sources. A dose of UVB was determined that was sufficient to stop cell growth but was not lethal. Fig. 1 shows the growth suppression of human melanocytes due to increasing doses of UV. UVB and UVC applied 12 h after seeding of the cells caused a dose-dependent delay in exponential growth. A dose of about 2000 J/m² UVB resulted in complete cessation of net growth for more than 3 days without reducing the viability below 85%. This is about one-half of the cumulative doses used to induce malignant melanomas in some animal models (23, 24).

UVB and TPA as Interacting but Largely Independent Modulators of RNA Abundance of Late Transcribed Target Genes. An experiment was designed to determine the proportion of transcripts that were regulated by overlapping transcriptional programs of UVB and TPA, as well as the proportion of genes regulated by TPA or UVB alone. To divide these transcripts between the primary response class, which is independent of new protein synthesis, and the secondary response class, which requires new protein synthesis, CX, an inhibitor of protein synthesis, was included as a third dissecting variable (25). Normal melanocytes from newborns can divide at least 15 times. These cells were used in the exponential growth phase after four to six cell divisions. At this time, they showed the dendricity and morphology of young, nonsenescent melanocytes. Cells received fresh media each day during days 1–6. On day seven, all cells received growth factor-free media for a period of 20 h. On day eight, cells were treated with UV/TPA/CX in all possible (2³ = 8) combinations: 0/0/0, 0/0/CX, 0/UV/0, TPA/0/0, TPA/UV/0, TPA/0/CX, 0/UV/CX, and UV/TPA/CX. In pilot experiments, the dose of 2000 J/m² UV B, as well as TPA and CX doses, was tested for nontoxicity, revealing at

[4] F. Mathieu-Daude, personal communication.

least 85% viable cells after 48 h of culture. Thirty-two nM TPA is within a broad dose range (10–150 nM) that uniformly activates the growth of human melanocytes, probably by stimulating different protein kinase C isoforms and modulating their cellular location (26, 27). To analyze the late effects of these modulators, cells were harvested 8 hours after treatment, a time point when, due to the fast kinetics of the TPA- and the UV response, early induced transcripts such as c-*jun*, c-*fos*, and c-*myc* have already passed their maximum expression (8, 27, 28).

Total RNAs prepared from all treatment groups were surveyed by RAP-PCR. To avoid misleading results due to PCR products that reflect differences in RNA quality and concentration rather than alterations in transcript abundance, each RNA was fingerprinted at three 2-fold serial dilutions. Only those differentially amplified products that were reproducibly present at all three RNA concentrations of a particular treatment group were considered to be derived from differentially expressed genes (20). Fig. 2 shows a representative RAP-PCR fingerprint. The order of loading (from left to right) corresponds to the codes used for OCRs in Table 1. The total number of visible PCR products was counted in the control lanes of all gels.

TPA	o	o	o	o	+	+	+	+
UV B	o	o	+	+	+	+	o	o
CX	o	+	o	+	+	o	+	o
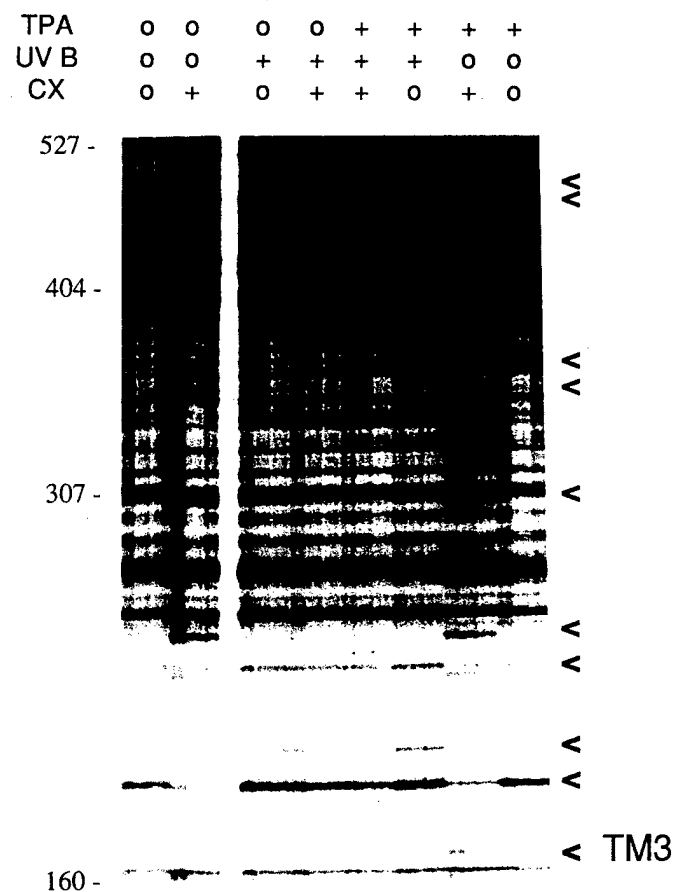
Fig. 2. RAP-PCR fingerprinting reveals divergent sets of UVB- and TPA-regulated transcripts at late time points. Cells were treated with 2000 J/m² UVB, 32 nM TPA, or 0.02 mg/ml CX in all possible (2³ = 8) combinations. RNAs of all treatment groups were harvested 8 h after treatment and were fingerprinted at three 2-fold dilutions (400, 200, and 100 ng of total RNA) using the arbitrary primers AP-11 (5'-AGGGGCACCA-3') and AP-1 (5'-GAGGGTGCCTT-3'). RAP-PCR products were loaded side-by-side, such that each set of three lanes contains the three RNA concentrations from each treatment group. The RAP-PCR products were electrophoresed on a 8 M urea/6% polyacrylamide gel. *TM3*, band from which we obtained a differentially regulated sequence that matches the human *TM3* gene. <, bands displaying differential amplification of a cDNA. *Left*, size of the amplified cDNAs.

Table 1 *Differential expression results from RAP-PCR analysis of melanocytes treated with all combinations of UVB, TPA, and CX*

		Treatments								No. of regulated transcripts observed[b]	
	TPA:	0	0	0	0	+	+	+	+		
	UV:	0	0	+	+	+	+	0	0		
OCR[a]	CX:	0	+	0	+	+	0	+	0		
ORC 02		0	+	0	0	0	0	+	0	51	a
ORC 08		0	−	0	0	0	0	−	0	8	
ORC 06		0	0	0	0	0	0	+	0	15	a
ORC 10		0	0	0	0	0	0	−	0	6	
ORC 07		0	0	0	0	0	0	+	+	13	a
ORC 21		0	0	0	0	0	0	−	−	1	
ORC 03		0	0	+	+	+	+	0	0	22	a
ORC 04		0	0	−	−	−	−	0	0	20	
ORC 22		0	0	+	+	+	+	−	0	1	a
ORC 18		0	0	−	−	−	−	+	0	2	
ORC 14		0	0	+	+	+	+	0	+	3	a
ORC 23		0	0	−	−	−	−	0	−	1	
ORC 11		0	0	+	+	+	+	+	+	6	a
ORC 19		0	0	−	−	−	−	−	−	2	
ORC 12		0	+	+	+	+	+	+	+	4	a
ORC 15		0	−	−	−	−	−	−	−	3	
ORC 05		0	+	0	0	0	0	+	+	17	b
ORC 09		0	0	0	0	0	0	0	+	8	
ORC 16		0	0	0	0	+	+	+	+	3	
ORC 17		0	+	0	0	0	0	0	0	3	
ORC 26		0	+	0	0	0	0	0	+	1	
ORC 24		0	+	0	+	+	0	+	0	1	
ORC 20		0	−	−	−	−	−	−	0	2	c
ORC 25		0	0	0	0	−	−	0	0	1	
ORC 27		0	−	0	0	0	−	−	−	1	
ORC 13		0	+	−	−	−	−	0	−	4	d
ORC 28		0	0	+	+	0	0	−	−	1	
ORC 29		0	0	0	0	−	+	+	+	1	
ORC 30		0	0	+	−	−	+	−	+	1	
ORC 31		0	−	−	−	−	−	−	+	1	
ORC 32		0	−	−	−	−	−	+	+	1	
ORC 33		0	−	0	0	0	0	0	+	1	
ORC 01		0	0	0	0	0	0	0	0	~1695	e

[a] OCRs were arbitrarily numbered according to size, with OCR 1 being the largest. +, −, 0 indicate increase, decrease, or no change in the intensity of the band in the RAP-PCR fingerprint, respectively.

[b] Two hundred five of approximately 1900 surveyed transcripts fell into 32 distinct OCRs, and ~1695 fell into a large unregulated class (OCR 01). The OCRs are grouped according to general features including: a, inverted response pairs; b, all positive responses; c, all negative responses; d, unpaired, mixed positive and negative responses; e, unresponsive to any of the three agents.

Using 11 arbitrary primers in 10 pairwise combinations, bands corresponding to about 1900 mRNAs were visible. Two hundred five of the transcripts, about 11% of all transcripts surveyed, displayed differential expression under the conditions tested. The 205 transcripts fell into 32 OCRs other than OCR 1 (no regulation, about 1695 of the observed transcripts). Because three treatments (UVB, TPA, and CX) can be used in eight possible combinations, the theoretical complexity of regulatory information amounts to 2187 (3^7) possible OCRs when considering only the direction of the response, *i.e.*, fragments showing down- or up-regulation, including new appearance and complete disappearance of a fragment, and no change. Thus, the total number of OCRs amounted to 1.5% of the theoretical number of 2187.

Table 1 shows all OCRs 8 h after the cells underwent treatments with the eight possible combinations of 2000 J/m² UVB, 32 nM TPA, and 0.02 mg/ml CX. Twenty OCRs (61% of total) contained only one to three regulated transcripts. Twelve OCRs contained between 4 and 51 transcripts. TPA increased the expression of 77 transcripts and decreased the expression of 24 transcripts. UVB up-regulated 38 transcripts and down-regulated 147 transcripts. TPA and UVB effects on differential gene expression overlapped at this time point. The overlap contained 20 transcripts (1.1% of all transcripts surveyed; OCRs 11, 12 15, 17, 19, 24, 25, and 30). However, of the 77

transcripts up-regulated by TPA, only 14 were also up-regulated by UVB. Sixty of the remaining 63 TPA up-regulated transcripts were dominantly down-regulated by UV.

OCR 4 (1.1%), which is not affected by TPA or CX, decreased in size with lower doses to 0.5% (at 1500J/m²) and 0.1% (at 750 J/m²) but remained large when nonphysiological UVC was used in doses (20–40 J/m²) comparable to those used in most previous studies of the UV response (2). On the other hand, many transcripts increased in abundance after UVB treatment (OCR 3, 22 transcripts, 1.2%). The size of this OCR remained unchanged when UVC was used in doses of 20 and 40 J/m².

In these experiments, all UVB-specific effects on RNA abundance, as well as the impact of UVB on TPA- and CX-induced effects on RNA abundance, were primary responses in that they occurred in the presence of CX (*e.g.*, OCRs 3 and 4). In contrast, 15 of the 77 TPA up-regulated messages were up-regulated only in the absence of CX (OCRs 9, 30, 14, 33, 31, 26), possibly indicating a requirement for intervening protein synthesis. OCR 5, with 17 transcripts (0.9%), indicated partially overlapping CX and TPA effects; these transcripts showed increased abundance after treatment with CX or TPA, and TPA/CX in combination. By far, the largest class was OCR 2 (51 transcripts, 2.7%), with transcripts that accumulate due to CX, many

A

TPA	o	o	o	o	+	+	+	+
UV B	o	o	+	+	+	+	o	o
CX	o	+	o	+	+	o	+	o

 <

TPA	o	o
UV B	o	+
CX	o	o

COXI   - 2.4
- 1.4

G3PDH

B

TPA	o	o	o	o	+	+	+	+
UV B	o	o	+	+	+	+	o	o
CX	o	+	o	+	+	o	+	o

 <

TPA	o	+	+
UV B	o	o	+
CX	o	+	+

TM3   ⌐ 2.4
- 1.4

G3PDH

C

TPA	o	o	o	o	+	+	+	+
UV B	o	o	+	+	+	+	o	o
CX	o	+	o	+	+	o	+	o

 <

TPA	o	+	o
UV B	o	o	+
CX	+	+	+

HSU46075   - 4.4

G3PDH

D

TPA	o	o	o	o	+	+	+	+
UV B	o	o	+	+	+	+	o	o
CX	o	+	o	+	+	o	+	o

 <

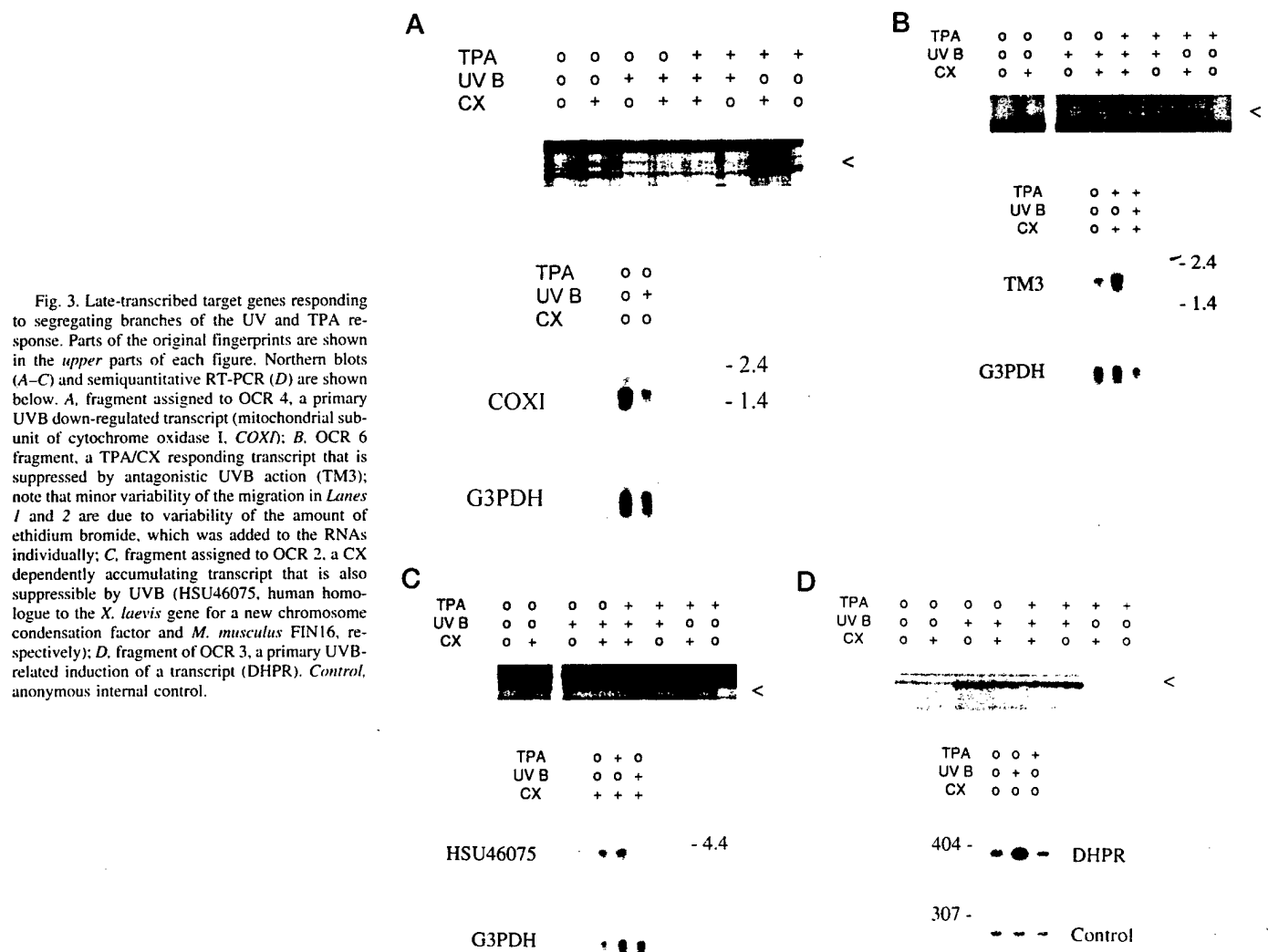|		
TPA	o	o	+
UV B	o	+	o
CX	o	o	o

404 -   DHPR

307 -   Control

Fig. 3. Late-transcribed target genes responding to segregating branches of the UV and TPA response. Parts of the original fingerprints are shown in the *upper* parts of each figure. Northern blots (*A–C*) and semiquantitative RT-PCR (*D*) are shown below. *A*, fragment assigned to OCR 4, a primary UVB down-regulated transcript (mitochondrial subunit of cytochrome oxidase I, *COXI*); *B*, OCR 6 fragment, a TPA/CX responding transcript that is suppressed by antagonistic UVB action (TM3); note that minor variability of the migration in *Lanes 1* and *2* are due to variability of the amount of ethidium bromide, which was added to the RNAs individually; *C*, fragment assigned to OCR 2, a CX dependently accumulating transcript that is also suppressible by UVB (HSU46075, human homologue to the *X. laevis* gene for a new chromosome condensation factor and *M. musculus* FIN16, respectively); *D*, fragment of OCR 3, a primary UVB-related induction of a transcript (DHPR). *Control*, anonymous internal control.

of which may result from inhibition of mRNA degradation (29). The accumulation of all transcripts in this class is blocked by UVB, possibly reflecting either global down-regulation of transcription or deactivation of the message stabilization program controlled by CX. Interestingly, however, global down-regulation of transcription cannot account for the behavior of OCR 8, suggesting that UV might interfere with CX action through a more pathway-specific mechanism. In 15 of the 23 OCRs where TPA has an impact on gene expression, UVB treatment has a dominant-negative effect over the TPA effect. Nine of the OCRs represent TPA primary responses that are overridden by UV. In all cases, the UV override signal is independent of CX. OCR 28 is a special case because it is the only observed situation where the UV effect might be overridden by TPA.

Further Characterization of Genes Representing the Main OCRs. Twenty transcripts were cloned and sequenced in this study, 4 of which showed very close homology or identity with sequences of known genes. These genes represent four of the major OCRs, including OCRs 6, 4, 3, and 2. To establish these four transcripts as markers of divergent sets of late-transcribed target genes, differential expression was confirmed by Northern blots and semiquantitative RT-PCR (Fig. 3). Differential expression of the other sixteen, putatively UVB/TPA/CX-regulated transcripts that did not match with known genes, has yet to be confirmed. Thus, the corresponding sequences were submitted as new melanocytic ESTs to the NCBI data base along with annotations regarding their regulation suggested by RAP-PCR exper-

iments (accession numbers U77311–22 up-regulated by UVB, U77323 up-regulated by TPA and UVB, U77324–6 up-regulated by TPA).

From OCR 4 (down-regulated by UVB regardless of the presence of CX or TPA), we characterized a differentially regulated transcript (about 5-fold down-regulation according to densitometric measurements) that entirely matched the sequence for the mitochondrial subunit of cytochrome oxidase I (GenBank entry HUMMTM1, between bases 2221 and 2480). RAP-PCR predicted a strong down-regulation by UV as a primary response. This was confirmed by Northern blots showing the expected transcript size of about 1600 bp (Fig. 3A). Because the mitochondrial genome is polycistronically organized, it is likely that all mitochondrial transcription was quickly and efficiently down-regulated within the first 8 h after UVB irradiation.

From OCR 6 (up-regulated by TPA in presence of CX and absence of UVB), a transcript was identified that completely aligned with nonmuscular human TM3 mRNA (GenBank entry HUMTRO, between bases 113 and 255), a molecule that was not previously known to be expressed in melanocytes. This is the first evidence for TPA/CX-dependent up-regulation of TM3 (about 15-fold up-regulation according to densitometric measurements) that is antagonized by UVB light in human melanocytes. Northern blots revealed a mRNA of about 2 kb, similar to the tropomyosin splicing variant found in human fibroblasts (30).

From OCR 2 (up-regulation by CX in the absence of UV B), a transcript was obtained that represented the first human homologue of a previously identified gene from *Xenopus laevis* (94 of 124 bases, 75%, are identical; GenBank entry XLU13674, between bases 3263 and 3386). This sequence was submitted to the NCBI data base (GenBank accession number U46075/HSU46075). Subsequently, a new entry in the data base showed a homologous *Mus musculus* sequence (GenBank accession U42385/MMU42385), which is inducible by fibroblast growth factor in mouse fibroblasts (31). Here, 107 of 124 bases, 87%, are identical with our sequence. This transcript is constitutively expressed in cultured melanocytes and accumulates when CX is present, and its 4-fold accumulation in the presence of CX is suppressed by simultaneous UVB irradiation. Northern analysis revealed a transcript of about 4.3 kb (Fig. 3C). This new family of genes encodes nucleotide binding proteins defined by the yeast protein Smc1p. One member of the family is thought to be involved in nuclear processes such as chromosome condensation during mitosis (32, 33).

From OCR 3 (UV up-regulated regardless of the presence of CX or TPA), a transcript was isolated that matched a clone of 18S rRNA (GenBank entry M16477). However, the sequence also shares homology with the *DHPR* gene (GenBank accession number M16447). DHPR maintains the supply of tetrahydrobiopterin, an essential cofactor of phenylalanine hydroxylase, an enzyme that controls melanin biosynthesis in human epidermis (34). UV induction (about 4-fold in our experiments according to densitometric measurements) of this gene is quite plausible from a physiological point of view. Because Northern blots were not sensitive enough in this case, we confirmed UV specific up-regulation by semiquantitative RT-PCR using primers, which produce a 402-bp product from the coding region of this gene (Fig. 3D).

The GenBank accession numbers were U77311–U77326 and U46075/HSU46075.

DISCUSSION

The response to UV and growth-promoting chemicals such as TPA in mammalian cells involves a limited repertoire of shared early-immediate transcribed genes that are activated by integrated signal transduction pathways (4, 8, 10, 35). Both pleiotropic agents lead to complex changes in transcriptional control, which may contribute to the adverse effects such as tumor initiation (UV) and promotion (UV and TPA; Refs. 26 and 28). Bearing in mind the caveats discussed previously (25), it is possible to assess the approximate scale of the TPA and UVB response by RAP-PCR. In our survey of about 1900 RAP-PCR products obtained from melanocytes treated with TPA, UVB, and CX in all possible combinations, 204 transcripts responded to TPA, UV, or both. Seventy-seven of these were up-regulated by TPA, but only 14 of these were also up-regulated by UVB. The overlap of 14 positively regulated genes 8 h after exposure to UVB and TPA may reflect persistent effects of common early-immediate responses (18). The limited extent of this overlap may be explained by the fact that TPA acts predominantly through the PKC pathway (36), whereas the pathways induced by UVB are more complex; the latter involves not only activation of SRC receptor tyrosine kinase but may also have effects on transcription resulting directly from DNA damage and DNA damage-dependent nuclear kinases (2).

The late responses to UVB, in contrast to TPA, were exclusively CX-resistant, primary responses. This is consistent with the widely accepted concept that posttranslational modification or release of preformed proteins by UV rather than new protein synthesis plays a major role in the UV response and makes it different from the response to growth factors and chemicals such as TPA, which seems to be more dependent on the biosynthesis of new proteins (2, 11). Accordingly, we detected up-regulation of 15 transcripts as a secondary response to TPA. That no secondary UVB response genes were observed may reflect a low abundance of the corresponding transcripts, causing them to be inefficiently sampled by RAP-PCR (37). Alternatively, categories that contain few regulated genes are likely to seem empty in a limited sample. For example, assuming unbiased sampling, a class representing 0.1% of all actively transcribed genes that can be sampled using this method (15 genes when assuming a total of 15,000 transcripts) have a 15% (0.999^{1900}) probability of not being sampled in a survey of 1900 transcripts. However, there is a 99.99% ($1-0.995^{1900}$) chance that at least one transcript will be sampled in a category containing 0.5% of all genes. Thus, we can be almost certain that if the seemingly empty categories have responsive transcripts, they represent less than 0.5% of all transcripts. Furthermore, more abundant cDNAs are partially excluded from amplification during later cycles of RAP-PCR due to reannealing, whereas rarer cDNAs amplify further ("C_0t effect"; Ref. 38). This partial abundance normalization and the fact that most fragments are independent of each other argue that our estimates of the size of each OCR are plausible (25, 39).

The identification of late-transcribed target genes of UVB pathways in melanocytes may provide new targets for tumor-preventive therapies as well as diagnostic and prognostic molecular markers for detecting UVB damage in precursor lesions associated with an increased risk of malignant transformation. Therefore, we identified marker transcripts of the main subsets of regulated transcripts observed that specifically respond to UVB or to UVB and one of the other experimental modulators, TPA and CX. Our search for genes responding to UVB revealed two new regulatory mechanisms that may contribute to the carcinogenic risk, which UVB inflicts on melanocytes over and above the direct mutagenic activity of UV due to pyrimidine dimer formation in nucleic acids. These new observations were down-regulation of mitochondrial transcription and of the tumor-suppressive transcript TM3.

The sampling of mRNAs deficient of a poly(A) tail, such as mitochondrial mRNAs or prokaryotic mRNAs, is a special feature of RAP-PCR when compared to poly-T anchored first-strand synthesis priming of differential display (19, 20). Therefore, in this study, a mitochondrial mRNA could be sampled, and we provide new evidence that UVB is able to efficiently down-regulate mitochondrial transcription in human melanocytes. This observation is consistent with the previously observed photodynamic action of UVA (UVA plus sensitizing chemicals such as psoralen; Ref. 40) and similar to the effects of some cytostatic drugs such as mofarotene (41). Poor mitochondrial function and accumulation of deletions in the mitochondrial genome has been described in nonmelanocytic skin cancer (42). Also, impaired mitochondrial transcription is associated with poor differentiation and clinical outcome in colon cancer (43). Thus, temporary or persistent mitochondrial malfunction may be a cocarcinogenic and tumor-promoting mechanism in a variety of human cancers (44) and could put the melanocyte at increased and prolonged mutagenic risk after UVB exposure due to lack of safe detoxification of oxygen radicals via the respiratory chain. This new aspect of indirect UVB mutagenicity may explain why repeated UVB exposure is much more harmful than single doses (2). The first hit of radical producing UV may leave the cells in a vulnerable status of impaired radical detoxification capacity so that another hit may cause aggravated DNA damage. Assuming that the mitochondrial function cannot be recovered completely by a subset of cells, these organelles could also conserve the *ad hoc* mutagenic effects of UVB due to a protracted or even life-long production of activated oxygen species in UVB-damaged cells. Further studies are needed to confirm the significance of

these mitochondrial alterations, as well as to develop possible preventive therapeutics that function as oxygen radical scavengers.

TM3 mRNA was not previously known to be expressed in human melanocytes. In addition to its marker function for late TPA-specific gene regulation, it has important implications for transformation and cancer (45). Several findings have suggested that the expression of RNAs coding for cytoskeletal proteins like TM3 are tumor protective; thus, expression returns if cells revert from malignant phenotype to benign and *vice versa* (45). Recently, it was demonstrated that increased expression of the untranslated region of this particular mRNA exerts the tumor-suppressing activity (45). Therefore, its up-regulation may reflect a transformation-protective response of normal melanocytes to strong tumor-promoting stimuli such as TPA. Our study shows that UVB light blocks this up-regulation of TM3 mRNA. This effect may reflect a new aspect of induction and promotion of melanocytic tumors by UV B. The disappearance of molecular markers such as TM3 mRNA transcript could become a new tool for diagnosis of melanocytic nevi that are at a labile, cancer-prone status (dysplasia), which is still a matter of controversy even among expert histopathologists (46). *In situ* studies of TM3 expression in this particular subset of melanocytic lesions are in progress.

We examined in detail one further transcript that constitutively accumulated in melanocytes in the presence of CX but was suppressible by UVB. The sequence of this transcript showed homology to already known genes, a basic fibroblast growth factor-inducible gene (*FIN16*) in mouse (31), as well as the *Xenopus laevis* gene for a coiled-coil protein that may act as a chromosome condensation protein during mitosis (32) and may also fulfill a yet unknown function in the cytoplasm throughout interphase (33). This family of genes, which is usually represented by transcripts of more than 4 kb (Fig. 3C), consists of members of an emerging family of nucleotide-binding proteins. The yeast protein Smc1p is the archetype in this family (33). The function of these proteins during interphase and where they localize in the cytoplasm are not yet known. Our sequence appears to represent the first human member of this family and was deposited in the NCBI GenBank data base (U46075/HSU46075).

A fourth transcript that we examined in detail codes for DHPR, which shows UVB-dependent up-regulation. This enzyme plays a pivotal role in maintaining the cellular pool of tetrahydrobiopterin, the major and essential cofactor of phenylalanin hydroxylase, which is the key enzyme of melanogenesis (34). This marker for UVB-specific effects most likely represents a physiological response in melanocytic cells, because implications for a close link to the carcinogenic risk are not apparent. However, as a paracrine factor, tetrahydrobiopterin also interferes with the nitric oxide production of vascular smooth muscle cells and may finally lead to vasodilatation in underlying tissues (47). Therefore, our observation might also contribute to further investigations that address secondary UVB-inducible phenomena, such as UVB erythema and sunburn, as well as vascularization in malignant melanomas. The *DHPR* gene belongs to a UV up-regulated OCR. Genes displaying this behavior have received much attention in the past due to their involvement in cancers. Some of the genes in this class, *e.g.*, ornithine decarboxylase, have become targets for the development of new cancer therapies (17). To date, there are more than 50 members in this class of genes, and it is still growing (2). Indeed, our estimates that about 1.9% of all transcribed genes in melanocytes are in this class indicate that it may be much larger and the effects of UV much more complex than previously thought. We sequenced a series of 12 more transcripts from this OCR that did not match any data base entries, which were submitted to GenBank as new melanocytic ESTs (GenBank accession numbers U77311-U77322).

Finally, there are several methods for the high throughput analysis of gene expression patterns, including RAP-PCR (20), differential display (19), SAGE (48), and differential hybridization to clones immobilized on membranes or glass slides (49) or oligos attached to glass slides (50). A number of laboratories are generating differential regulation data, and methods for analyzing the patterns that emerge from this data are evolving (25). Gene expression pattern analysis may reveal correlations between different genes and regulatory phenomena that could not be easily achieved by examining genes one at a time.

One type of pattern can be seen in Table 1, where many OCRs display inverted response behavior. For example, treatment by UV results in a positive response in OCR 3 and a negative response in OCR 4, possibly reflecting the simple fact that UV up-regulates some genes while down-regulating others. Interestingly, however, in paired OCRs 2 and 8, the effects of CX are blocked by UV whether the CX effect is positive (OCR 2) or negative (OCR 8). A similar comparison can be made for OCRs 6 and 10, where both positive and negative TPA responses, which are observed only in the presence of CX, are blocked by UV. OCRs 14 and 23 display inverted responses that reflect the ability of CX to block the effects of TPA, whether the effect is positive (OCR 14) or negative (OCR 23), and both positive and negative effects of TPA are blocked by UV in OCRs 7 and 21. Also in keeping with the theme of inverse regulatory classes, genes in OCR 12 respond to any of the three agents by up-regulation, and this class is mirrored by OCR 15, which is down-regulated by any of the three agents. Likewise, OCR 11 is up-regulated by either UV or TPA, and this class is mirrored by OCR 19, wherein all genes are down-regulated by either UV or TPA.

Considering that over 2000 response patterns are possible, it is surprising that so many genes fall into classes that constitute inverted response pairs. Genes that fall into the same OCRs are more likely to share regulatory features than genes from different classes. In addition, however, most genes in this study (78%) fell into inverted response pairs, suggesting that both positive and negative effects are often mediated by common signal transduction pathways, perhaps often by the same transcription factor. The alternative hypothesis, that the effects of the different agents are mediated by nonoverlapping signal transduction pathways that converge only at the promoter, would lead to a collection of OCRs with few inverted pairs, because each OCR would be an independent combination drawn from the 2187 possibilities. These principles must be explored in greater detail, but if correct, RNA abundance comparisons by RAP or any other method, under various treatment scenarios, may provide a means by which genes can be grouped according to common regulatory features. This would greatly aid in the design of directed experiments to identify the signal transduction subroutines responsible for controlling different classes of genes.

In conclusion, our data indicate global and sustained disruption of transcriptional control 8 h after a growth-arresting dose of UVB in human melanocytes. Also at this late time point, an overlapping set of genes was found being regulated by both UVB and TPA. However, our data also show that the pathways involved, which initially have intimate links, are ultimately able to induce divergent programs that reflect divergent and agent-specific phenotypes in melanocytic cells, *i.e.*, growth arrest by UVB and growth stimulation by TPA. The molecular markers identified in this study may facilitate further studies that address UV-promoted melanocytic transformation and progression. Our findings that UVB can alter mitochondrial functions and block the TPA-dependent up-regulation of the tumor-suppressive TM3 transcript underscores the concept that UV induces a partially transformed phenotype (11). Further studies are needed to determine whether mitochondrial impairment and consequent oxygen radical formation due to lesions caused by UV are responsible for the observation that single

UV exposures can result in the subsequent oncogene-driven accumulation of mutations in subsequent generations of cells (51), and whether this is the basic mechanism behind the onset of malignant melanoma many years after a few sunburns.

REFERENCES

1. Rigel, D., Friedman, R., and Kopf, A. The incidence of malignant melanoma in the United States: issues as we approach the 21st century. J. Am. Acad. Dermatol., 34: 839–847, 1996.
2. Friedberg, E., Walker, G., and Siede, W. Regulatory responses to DNA-damaging agents in eukaryotic cells. In: E. Friedberg (ed.), Regulatory Responses to DNA-damaging Agents in Eukaryotic Cells, pp. 595–631. Washington: American Society for Microbiology (ASM Press), 1995.
3. Little, J., and Mount, D. The SOS regulatory system of Escherichia coli. Cell, 29: 111–122, 1982.
4. Devary, Y., Gottlieb, R., Smeal, T., and Karin, M. The mammalian ultraviolet response is triggered by activation of Src tyrosine kinases. Cell, 71: 1081–1091, 1992.
5. Donawho, C., Wolf, P., and Kripke, M. Enhanced development of murine melanoma in UV-irradiated skin: UV dose response, waveband dependence, and relation to inflammation. Melanoma Res., 4: 93–100, 1994.
6. Radler-Pohl, A., Sachsenmaier, C., Gebel, S., Auer, H., Bruder, J., Rapp, U., Angel, P., Rahmsdorf, H., and Herrlich, P. UV-induced activation of AP-1 involves obligatory extranuclear steps including Raf-1 kinase. EMBO J., 12: 1005–1012, 1993.
7. Davis, R. MAPKs: new JNK expands the group. Trends Biochem. Sci., 19: 470–473, 1994.
8. Devary, Y., Gottlieb, R., Lau, L., and Karin, M. Rapid and preferential activation of the c-jun gene during the mammalian UV response. Mol. Cell. Biol., 11: 2804–2811, 1991.
9. Radler-Pohl, A., Gebel, S., Sachsenmaier, C., Konig, H., Kramer, M., Oehler, T., Streile, M., Ponta, H., Rapp, U., and Rahmsdorf, H. The activation and activity control of AP-1 (fos/jun). Ann. NY Acad. Sci., 684: 127–148, 1993.
10. Devary, Y., Rosette, C., DiDonato, J., and Karin, M. NF-κ B activation by ultraviolet light not dependent on a nuclear signal. Science (Washington DC), 261: 1442–1445, 1993.
11. Herrlich, P., Ponta, H., and Rahmsdorf, H. DNA damage-induced gene expression: signal transduction and relation to growth factor signaling. Rev. Physiol. Biochem. Pharmacol., 119: 187–223, 1992.
12. Leonardo, A., Linke, S., Clarkin, K., and Wahl, G. DNA damage triggers a prolonged p53 dependent G1 arrest and a long-term induction of CIP1 in normal human fibroblasts. Genes Dev., 8: 2540–2551, 1994.
13. Kroumpouzos, G., Eberle, J., Garbe, C., and Orfanos, C. p53 mutation and c-fos overexpression are associated with detection of the antigen VLA-2 in human melanoma cell lines. Pigment Cell Res., 7: 348–353, 1994.
14. Huang, L., Clarkin, K., and Wahl, G. Sensitivity and selectivity of the DNA damage sensor responsible for activating p53-dependent G1 arrest. Proc. Natl. Acad. Sci. USA, 93: 4827–4832, 1996.
15. Lazaris, A., Theodoropoulos, G., Aroni, K., Saetta, A., and Davaris, P. Immunohistochemical expression of c-myc oncogene, heat shock protein 70 and HLA-DR molecules in malignant cutaneous melanoma. Virchows Arch., 426: 461–467, 1995.
16. Muijen, G., van Danen, E., de Vries, T., Quax, P., Verheijen, J., and Ruiter, D. Properties of metastasizing and nonmetastasizing human melanoma cells. Recent Res. Cancer Res., 139: 105–122, 1995.
17. Porter, C., Ganis, B., Rustum, Y., Wrzosek, C., Kramer, D., and Bergeron, R. Collateral sensitivity of human melanoma multidrug-resistant variants to the polyamine analogue, N^1,N^{11}-diethylnorspermine. Cancer Res., 54: 4517–4524, 1994.
18. Whitmarsh, A., Shore, P., Sharrocks, A., and Davis, R. Integration of MAP kinase signal transduction pathways at the serum response element. Science (Washington DC), 269: 403–407, 1995.
19. Liang, P., and Pardee, A. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science (Washington DC), 257: 967–971, 1992.
20. Welsh, J., Chada, K., Dalal, S. S., Cheng, R., Ralph, D., and McClelland, M. Arbitrarily primed PCR fingerprinting of RNA. Nucleic Acids Res., 20: 4965–4970, 1992.
21. Mathieu Daude, F., Cheng, R., Welsh, J., and McClelland, M. Screening of differentially amplified cDNA products from RNA arbitrarily primed PCR fingerprints using single stranded conformation polymorphism (SSCP) gels. Nucleic Acids Res., 24: 1504–1507, 1996.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1989.
23. Klein-Szanto, A., Silvers, W., and Mintz, B. Ultraviolet radiation-induced malignant skin melanoma in melanoma-susceptible transgenic mice. Cancer Res., 54: 4569–4572, 1994.
24. Robinson, E., VandeBerg, J., Hubbard, G., and Dooley, T. Malignant melanoma in ultraviolet irradiated laboratory opossums: initiation in suckling young metastasis in adults and xenograft behavior in nude mice. Cancer Res., 54: 5986–5991, 1994.
25. McClelland, M., Ralph, D., Cheng, R., and Welsh, J. Interactions among regulators of RNA abundance characterized using RNA fingerprinting by arbitrarily primed PCR. Nucleic Acids Res., 22: 4419–4431, 1994.
26. Bennett, D. Mechanisms of differentiation in melanoma cells and melanocytes. Environ. Health Perspect., 80: 49–59, 1989.
27. Rahmsdorf, H., and Herrlich, P. Regulation of gene expression by tumor promotors. Pharmacol. Ther., 48: 157–188, 1990.
28. Medrano, E., Im, S., Yang, F., and Abdel-Malek, Z. Ultraviolet B light induces G_1 arrest in human melanocytes by prolonged inhibition of retinoblastoma protein phosphorylation associated with long-term expression of the $p21^{Waf-1/SDI-1/Cip-1}$ protein. Cancer Res., 55: 4047–4052, 1995.
29. Williams, D., Sensel, M., McTigue, M., and Binder, R. Hormonal and developmental regulation of mRNA turnover. In: J. Belasco and G. Brawerman (eds.), Hormonal and Developmental Regulation of mRNA Turnover, pp. 161–91. San Diego: Academic Press, Inc., 1993.
30. Novy, R. E., Lin, J. L., Lin, C. S., and Lin, J. J. Human fibroblast tropomyosin isoforms: characterization of cDNA clones and analysis of tropomyosin expression in human tissues and in normal and transformed cells. Cell Motil. Cytoskeleton, 25: 267–281, 1993.
31. Guthridge, M., Seldin, M., and Basilico, C. Induction of expression of growth-related genes by FGF-4 in mouse fibroblasts. Oncogene, 12: 1267–1278, 1996.
32. Hirano, T., and Mitchison, T. A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. Cell, 79: 449–458, 1994.
33. Saitoh, N., Goldberg, I., Wood, E., and Earshaw, W. SCII: an abundant chromosome scaffold protein is a member of a family of putative ATPases with an unusual predicted tertiary structure. J. Cell Biol., 127: 303–318, 1994.
34. Schallreuter, K., Wood, J., Pittelkow, M., Gutlich, M., Lemke, K., Rodl, W., Swanson, N., Hitzemann, K., and Ziegler, I. Regulation of melanin biosynthesis in human epidermis by tetrahydrobiopterin. Science (Washington DC), 263: 1444–1446, 1994.
35. Rosen, C., Gajic, D., Jia, Q., and Drucker, D. Ultraviolet B radiation induction of ornithine decarboxylase gene expression in mouse epidermis. Biochem. J., 12: 565–568, 1990.
36. Valyi-Nagy, I., and Herlyn, M. Regulation of growth and phenotype of normal human melanocytes in culture. In: L. Nathanson (ed.), Regulation of Growth and Phenotype of Normal Human Melanocytes in Culture, pp. 85–101. Boston: Kluwer Academic Publishers, 1991.
37. Ralph, D., McClelland, M., and Welsh, J. RNA fingerprinting using arbitrarily primed PCR identifies differentially regulated RNAs in mink lung (MvILu) cells growth arrested by transforming growth factor β 1. Proc. Natl. Acad. Sci. USA, 90: 10710–10714, 1993.
38. Mathieu-Daude, F., Welsh, J., Vogt, T., and McClelland, M. DNA rehybridization during PCR: the "cot effect" and its consequences. Nucleic Acids Res., 24: 2080–2086, 1996.
39. McClelland, M., Mathieu-Daude, F., and Welsh, J. RNA fingerprinting and differential display using arbitrarily primed PCR. Trends Genet., 11: 242–246, 1995.
40. Salet, C., and Moreno, G. Photodynamic action increases leakage of the mitochondrial electron transport. Int. J. Radiat. Biol., 67: 477–480, 1995.
41. Uchida, T., Inagaki, N., Furuichi, Y., and Eliason, J. Down-regulation of mitochondrial gene expression by the anti-tumor arotinoid mofarotene (Ro 40-8757). Int. J. Cancer, 58: 891–897, 1994.
42. Pang, C., Lee, H., Yang, J., and Wei, Y. H. Human skin mitochondrial DNA deletions associated with light exposure. Arch. Biochem. Biophys., 312: 534–538, 1994.
43. Heerdt, B., Halsey, H., Lipkin, M., and Augenlicht, L. Expression of mitochondrial cytochrome c oxidase in human colonic cell differentiation, transformation, and risk for colonic cancer. Cancer Res., 50: 1596–1600, 1990.
44. Cotton, D. Ageing, cancer and mitochondrial deterioration. Lancet, 341: 281–282, 1993.
45. Rastinejad, F., Conboy, M., Rando, T., and Blau, H. Tumor suppression by RNA from the 3' untranslated region of α-tropomyosin. Cell, 75: 1107–1117, 1993.
46. Stolz, W., Vogt, T., Landthaler, M., Hempfer, S., Bingler, P., and Abmayr, W. Differentiation between malignant melanoma and benign melanocytic nevi by computerized DNA cytometry of imprint specimens. J. Cutaneous Pathol., 21: 7–15, 1994.
47. Schaffner, A., Blau, N., Schneemann, M., Steurer, J., Edgell, C., and Schoedon, G. Tetrahydrobiopterin as another EDRF in man. Biochem. Biophys. Res. Commun., 205: 516–523, 1994.
48. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. Serial analysis of gene expression. Science (Washington DC), 270: 484–487, 1995.
49. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science (Washington DC), 270: 467–470, 1995.
50. Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. C., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S., and Fodor, S. P. Accessing genetic information with high-density DNA arrays. Science (Washington DC), 274: 610–614, 1996.
51. Little, J., Gorgojo, L., and Ventrovs, H. Delayed appearance of lethal and specific gene mutations in irradiated mammalian cells. Int. J. Radiat. Oncol. Biol. Phys., 19: 1425–1430, 1990.

Non-stoichiometric reduced complexity probes for cDNA arrays

Thomas Trenkle, John Welsh*, Barbara Jung, Francoise Mathieu-Daude and Michael McClelland

Sidney Kimmel Cancer Center, 10835 Altman Row, San Diego, CA 92121, USA

ABSTRACT

A method is presented in which the reduced complexity and non-stoichiometric amplification intrinsic to RNA arbitrarily primed PCR fingerprinting (RAP-PCR) is used to advantage to generate probes for differential screening of cDNA arrays. RAP-PCR fingerprints were converted to probes for human cDNA clones arrayed as *Escherichia coli* colonies on nylon membranes. Each array contained 18 432 cDNA clones from the IMAGE consortium. Hybridization to ~1000 cDNA clones was detected using each RAP-PCR probe. Different RAP-PCR fingerprints gave hybridization patterns having very little overlap (<3%) with each other or with hybridization patterns from total cDNA probes. Consequently, repeated application of RAP-PCR probes allows a greater fraction of the message population to be screened on this type of array than can be achieved with a radiolabeled total cDNA probe. This method was applied to RNA from HaCaT keratinocytes treated with epidermal growth factor. Two RAP-PCR probes detected hybridization to 2000 clones, from which 22 candidate differentially expressed genes were observed. Differential expression was tested for 15 of these clones using RT–PCR and 13 were confirmed. The use of this cDNA array to analyze RAP-PCR fingerprints allowed for an increase in detection of 10–20-fold over the conventional denaturing polyacrylamide gel approach to RAP-PCR or differential display. Throughput is vastly improved by the reduction in cloning and sequencing afforded by the use of arrays. Also, repeated cloning and sequencing of the same gene or of genes already known to be regulated in the system of interest is minimized. The procedure we describe uses inexpensive arrays of plasmid clones spotted as *E.coli* colonies to detect differential expression, but these reduced complexity probes should also prove useful on arrays of PCR-amplified fragments and on oligonucleotide chips. Genes observed in this manuscript: H11520, U35048, R48633, H28735, M13918, H12999, H05639, X79781, M31627, H23972, AB000712, R75916, U66894, AF067817.

INTRODUCTION

Arrays of cDNA clones or oligonucleotides affixed to a solid support can capture labeled homologous cDNA from solution and, thereby, measure the differential expression of many genes in parallel. However, a total labeled cDNA probe from a mammalian cell typically has a complexity of >30 000 000 bases, which complicates attempts to detect differential expression among the rarer mRNAs using differential hybridization. Recent advances in the use of fluorescence and confocal microscopy have led to improvements in the sensitivity and dynamic range of differential hybridization methods and the detection of transcripts at a sensitivity approaching 1/500 000 (1,2 and references therein).

Despite these improvements, several of these methods are currently too expensive for the average molecular biology laboratory to implement. On the other hand, arrays of *Escherichia coli* colonies containing tens of thousands of sequenced ESTs are available for differential screening and are quite inexpensive. The standard method for differential screening, which typically uses probes derived from reverse transcription of total message and autoradiography or phosphorimaging, can give impressive results (3). However, the method is limited to the most abundant messages; only these abundant transcripts are represented highly enough to yield effective probes with a sensitivity of perhaps 1/15 000 (4). Here we show that differential screening of arrays of plasmids in colonies can be improved greatly by reducing the complexity of the probe and by systematically increasing the contribution of rarer mRNAs to the probe. In this way, differential screening using these arrays is not confined to only the most abundant mRNAs.

One way to construct a probe having reduced complexity and increased representation of rare messages is to use RAP-PCR fingerprinting, which samples a reproducible subset of the message population based on the best matches with arbitrary primers (5,6). In a typical RAP-PCR fingerprint, ~50–100 cDNA fragments per lane are visible on a polyacrylamide gel, including products from relatively rare mRNAs that happen to have among the best matches with the arbitrary primers. If only 100 cDNA clones could be detected in an array by each probe, then hybridization to arrays would be inefficient. However, RAP-PCR fingerprints contain many products that are too rare to visualize by autoradiography of a polyacrylamide gel. Nonetheless, these

*To whom correspondence should be addressed. Tel: +1 619 450 5990; Fax: +1 619 550 3998; Email: jwelsh@skcc.org

rarer products are reproducible and of sufficient abundance to serve as probe for arrays when labeled at high specific activity.

The experiments presented here show that a single probe derived from RAP-PCR can detect ~1000 cDNAs on an array containing ~18 432 EST clones, a 10–20-fold improvement over the performance of fingerprints displayed on denaturing polyacrylamide gels. In addition, when a differentially regulated gene is detected on a cDNA array, a clone representing the transcript is immediately available and often sequence information for the clone is also available. Furthermore, the clones are usually much longer than the usual RAP-PCR product. In contrast, the standard approaches to RNA fingerprinting require that the product be gel purified and sequenced before verification of differential expression can be performed.

In this report, we show that expression differences that can be seen in a standard RAP-PCR fingerprint can also be detected using fingerprints as differential screening probes against arrays. We further show that differentially amplified RAP-PCR products that are below the detection capabilities of the standard denaturing polyacrylamide gel and autoradiography methods can be detected using hybridization to cDNA arrays.

MATERIALS AND METHODS

RNA preparation

The immortal human keratinocyte cell line HaCaT (7) was grown to confluence and maintained at confluence for 2 days. The medium (Dulbecco's modified Eagle's medium + 10% fetal bovine serum + penicillin/streptomycin) was changed 1 day prior to experiments. Epidermal growth factor (EGF) (Gibco-BRL) was added at 20 ng/ml or transforming growth factor TGF-β (R&D systems, Minneapolis, MN) was added at 5 ng/ml. Treated and untreated cells were harvested after 4 h by scraping the Petri dishes in the presence of lysis buffer (Qiagen, Chatsworth, CA) and homogenized through Qiashredder columns. On average, 7×10^6 cells (confluent growth in a 100 mm diameter Petri dish) yielded 40 µg total RNA from the RNeasy total RNA purification kit (Qiagen). RNA, in 20 mM Tris, 10 mM MgCl$_2$ buffer, was incubated with 0.08 U/µl RNase-free DNase and 0.32 U/µl RNase inhibitor (both from Boehringer Mannheim Biochemicals, Indianapolis, IN) for 40 min at 37°C and cleaned again using the RNeasy kit. This step is important because small amounts of genomic DNA can contribute to the fingerprints. RNA quantity was measured by spectrophotometry and RNA samples were adjusted to 400 ng/µl in water. They were checked for quality and concentration by agarose gel electrophoresis and stored at –20°C.

RNA fingerprinting

RAP-PCR was performed using standard protocols (8,9). Reverse transcription was performed on total RNA using four concentrations per sample (1000, 500, 250 and 125 ng/ reaction) and an oligo(dT) primer (15mer) (Genosys Biotechnologies, The Woodlands, TX). RNA (5 µl) was mixed with 5 µl RT mixture for a 10 µl final reaction containing 50 mM Tris, pH 8.3, 75 mM KCl, 3 mM MgCl$_2$, 20 mM DTT, 0.2 mM each dNTP, 0.5 µM primer and 20 U MuLV reverse transcriptase (Promega, Madison, WI). RNA samples were checked for DNA contaminants by including a reverse transcriptase-free control in initial RAP-PCR experiments. The reaction was performed at 37°C for 1 h (after a 5 min ramp from 25 to 37°C), the enzyme was inactivated by heating the samples at 94°C for 5 min and the newly synthesized cDNA was diluted 4-fold in water.

PCR was performed after the addition of a pair of two different 10mer or 11mer oligonucleotide primers of arbitrary sequence: pair A, GP14 (GTAGCCCAGC) and GP16 (GCCACCCAGA); pair B, Nuc1+ (ACGAAGAAGAAGAG) and OPN24 (AGGG-GCACCA). In general, there are no particular constraints on the primers except that they contain at least a few C or G bases, that the 3'-ends are not complementary with themselves or the other primer in the reaction, to avoid primer dimers, and that primer sets are chosen that are different in sequence so that the same parts of mRNA are not amplified in different fingerprints.

Diluted cDNAs (10 µl) were mixed with the same volume of 2× PCR mixture containing 20 mM Tris, pH 8.3, 20 mM KCl, 6.25 mM MgCl$_2$, 0.35 mM each dNTP, 2 µM each oligonucleotide primer, 2 µCi [α-^{32}P]dCTP (ICN, Irvine, CA) and 5 U AmpliTaq® DNA polymerase Stoffel fragment (Perkin-Elmer Cetus, Norwalk, CT) for a 20 µl final reaction. Thermocycling was performed using 35 cycles of 94°C for 1 min, 35°C for 1 min and 72°C for 2 min.

An aliquot of the amplification products (3.5 µl) was mixed with 9 µl formamide dye solution, denatured at 85°C for 4 min and chilled on ice. A sample of 2.4 µl was loaded onto a 5% polyacrylamide, 43% urea gel, prepared with 1× TBE buffer. The PCR products resulting from the four different concentrations of the same RNA template were loaded side by side on the gel (Fig. 1).

Electrophoresis was performed at 1700 V or at a constant power of 50–70 W until the xylene cyanol tracking dye reached the bottom of the gel (~4 h). The gel was dried under vacuum and placed on Kodak BioMax X-Ray film for 16–48 h.

Labeling of RAP-PCR products for use as probes against cDNA arrays

Up to 10 µg PCR product from RAP-PCR can be purified using a QIAquick PCR Purification Kit (Qiagen, Chatsworth, CA), which removes unincorporated bases, primers and primer dimers <40 bp. The DNA was recovered in 50 µl 10 mM Tris, pH 8.3.

Random primed synthesis with incorporation of [α-^{32}P]dCTP was performed using a standard protocol. Ten percent of the recovered fingerprint DNA (typically ~100 ng in 5 µl) was combined with 3 µg random hexamer oligonucleotide primer and 0.3 µg each of the fingerprint primers in a total volume of 14 µl, boiled for 3 min and then placed on ice.

The hexamer/primer/DNA mix was mixed with 11 µl reaction mix to yield a 25 µl reaction containing 0.05 mM three dNTPs (minus dCTP), 50 µCi [α-^{32}P]dCTP (3000 Ci/mmol, 5 µl), 1× Klenow fragment buffer (50 mM Tris–HCl, 10 mM MgCl$_2$, 50 mM NaCl, pH 8.0, and 4 U Klenow fragment; Gibco-BRL Life Technologies, Gaithersburg, MD). The reaction was performed at room temperature for 4 h. For maximum probe length the reaction was chased by adding 1 µl 1.25 mM dCTP and incubated for 15 min at 25°C, then for an additional 15 min at 37°C. The unincorporated nucleotides, hexamers and primers were removed with the Qiagen Nucleotide Removal Kit and the purified products were eluted using two aliquots of 140 µl 10 mM Tris, pH 8.3.

Labeling of poly(A)+ mRNA and genomic DNA for use in arrays

Poly(A)+-selected mRNA and genomic DNA were labeled using random hexamers. Genomic DNA (150 ng) was labeled using the

same protocol used for labeling the RAP-PCR products. Poly(A)$^+$ mRNA (1 μg) and 9 μg random hexamer in a volume of 27 μl were incubated at 70°C for 2 min and chilled on ice. The RNA/hexamer mix was mixed with 23 μl master mix containing 10 μl 5× AMV reaction buffer (250 mM Tris–HCl, pH 8.5, 40 mM MgCl$_2$, 150 mM KCl and 5 mM DTT), 1 μl 33 mM each of three dNTPs (minus dCTP), 2 μl AMV reverse transcriptase (Boehringer Mannheim, Indianapolis, IN) and 10 μl [α-^{32}P]dCTP (3000 Ci/mmol) in a final volume of 50 μl. The reaction was incubated at room temperature for 15 min, ramped for 1 h to 47°C, held at 47°C for 1 h and chased with 1 μl 33 mM dCTP for another 30 min at 47°C. The labeled products were purified as described above.

Hybridization to the array

When radioactivity is used to label the probe, four membranes are needed, one membrane for each of two concentrations of RNA for each of the two RNA samples to be compared. If two color fluorescence were to be used, then two arrays would be needed, one for each of the two concentrations of starting RNA, because probes from the two RNA samples can be mixed. Here we present a protocol for radiolabeled probes.

Prewash of cDNA filters. The cDNA filters (Genome Systems, St Louis, MO) were washed in three changes of 2× SSC, 0.1% SDS in a horizontally shaken flat bottomed container to reduce the residual bacterial debris. The first wash was carried out in 500 ml for 10 min at room temperature. The second and third washes were carried out in 1 l prewarmed (55°C) prewash solution for 10 min each.

Prehybridization. The filters were transferred to roller bottles and prehybridized in 60 ml prewarmed (42°C) prehybridization solution containing 6× SSC, 5× Denhardt's reagent, 0.5% SDS, 100 μg/ml fragmented denatured salmon sperm and 50% formamide for 1–2 h at 42°C in an oven.

Hybridization. The prehybridization solution was exchanged with 7 ml prewarmed (42°C) hybridization solution containing 6× SSC, 0.5% SDS, 100 μg/ml fragmented denatured salmon sperm and 50% formamide. To decrease the background hybridization due to repeats (e.g. Alu and Line elements), sheared human genomic DNA was denatured in a boiling water bath for 10 min and immediately added to the hybridization solution to a final concentration of 10 μg/ml. An aliquot of 10 ng/ml poly(dA) can be added to block oligo(dT) stretches in the radiolabeled probe. Simultaneously, the labeled probe, in a total volume of 280 μl, was denatured in a boiling water bath for 4 min and immediately added to the hybridization solution. The hybridization was carried out at 42°C for 2–48 h (typically 18 h) in large roller bottles.

Wash. For the washes the incubator oven temperature was set to 68°C. The hybridization solution was poured off and the membrane was washed twice with 50 ml 2× SSC, 0.1% SDS at room temperature (RT) for 5 min. The wash solution was then replaced with 100 ml 0.1× SSC, 0.1% SDS (RT) and incubated for 10 min. Further washes were performed in 100 ml 0.1× SSC, 0.1% SDS at 55–68°C for 40 min in the roller bottles, followed by washing in 1 l for 20 min with gentle agitation in a horizontal shaker. The filters were transferred back to the roller bottles containing 100 ml prewarmed (55–68°C) 0.1× SSC, 0.1% SDS

and incubated for 1 h. The final wash solution was removed and the filters were briefly rinsed in 2× SSC at room temperature.

After washing, the membranes were blotted with 3MM paper, wrapped in Saran wrap while moist and placed against X-ray film. The membranes were usually sufficiently radioactive that a 1 day exposure with a screen will reveal the top 1000 products on an array of 18 432 bacterial colonies carrying EST clones. Weaker probes or fainter hybridization events can be seen using an intensifying screen at –70°C for a few days. Also, membranes may be read using a phosphorimager or using a fluorescence scanner when fluorescent probes are used.

Confirmation of differential expression using low stringency RT–PCR

The first level of confirmation is the use of two RNA concentrations per sample. Only those hybridization events that seem to indicate differential expression at both RNA concentrations in both RNA samples can be relied upon.

More than 70% of the IMAGE consortium clones have single pass sequence reads from the 5'- or 3'-end or both deposited in the GenBank database. In cases where there is no prior sequence information available, the clones can be ordered from Genome Systems and sequenced. Sequences were used to derive PCR primers of 18–25 bases in length using MacVector 6.0 (Oxford Molecular Group, Oxford, UK). Generally, primers were chosen to generate PCR products of 50–250 bp and have melting temperatures of at least 60°C.

Reverse transcription was performed under the same conditions as in the RAP-PCR protocol (above), using an oligo(dT) primer or a mixture of random 9mer primers (Genosys). The PCR reaction was performed using two specific primers (18–25mer). The PCR conditions were the same as in the RAP-PCR fingerprint protocol but 1.5 μM each primer was used. The following low stringency thermal profile was used: 94°C for 40 s, 47°C for 40 s and 72°C for 1 min, for 19, 22 and 25 cycles. The reactions were carried out in three sets of tubes at different cycle numbers because the abundance of the transcripts, the performance of the primer pairs and the amplifiability of the PCR products can vary. PCR products were run under the same conditions as above on a 5% polyacrylamide, 43% urea gel. The gel was dried and exposed for 18–72 h. Invariance among the other arbitrary products in the fingerprint was used as an internal control to indicate the reliability of the relative quantitation. Primer pairs (Genosys) used for confirmation of differential expression were as follows: GenBank accession no. H11520 (90 nt product), (A) AATGAG-GGGGACAAATGGGAAGC, (B) GGAGAGCCCTTCCTCA-GACATGAAG; TSC-22 gene (U35048, H11073, H11161) (179 nt product), (A) TGACAAAATGGTGACAGGTAGCTGG, (B) AA-GTCCACACCTCCTCAGACAGCC; R48633 (178 nt product), (A) CCCAGACACCCAAACAGCCGTG, (B) TGGAGCAGCC-GTGTGTGCTG. Figure 3 was assembled using Adobe PhotoShop.

RESULTS AND DISCUSSION

Choice of array

Arrays containing cDNA clones are available on nylon membranes from a variety of suppliers, including Research Genetics (www.resgen.com), Genome Systems (www.genomesystems.com) and the German Human Genome Project (www.rzpd.de). These arrays include clones from various human tissues, stages of

development and disease states. Arrays of mouse and yeast sequences are also available. At present, there are two types of arrays available on nylon membranes. One type contains 18 432 *E.coli* colonies, each carrying a different IMAGE consortium EST plasmid (www-bio.llnl.gov/bbrp/image/image.html), spotted twice on a 22 × 22 cm membrane (available from Genome Systems). The second type contains >5000 PCR products from selected IMAGE clones amplified using vector primers, available from Research Genetics. To date, an array of PCR products is available for every yeast ORF and for a subset of human ESTs. One can expect a dramatic increase in the number of available arrays, organisms and accompanying sequence information.

We chose the Genome Systems arrays, which contain by far the largest number of ESTs per unit cost. However, each spotted EST is associated with *E.coli* genomic DNA from the host, in contrast to PCR product arrays and oligonucleotide arrays which are free of other DNAs. Thus, the clone arrays should have the highest background among the current choices and represent the greatest challenge for the probes we developed.

RNA fingerprinting for probe preparation

RAP-PCR amplifications were performed to look for differential gene expression in keratinocytes (HaCaT) (7) when treated with EGF or TGF-β for 4 h. Using RAP-PCR, ~1% of the genes in normal or immortal keratinocytes responded to EGF and fewer responded to TGF-β in this time frame (data not shown). Two fingerprints were chosen for hybridization to cDNA arrays. Figure 1A and B shows RAP-PCR fingerprints of RNA from confluent keratinocytes treated with TGF-β or EGF, using multiple RNA concentrations and two sets of arbitrarily chosen primers. Primarily, the untreated control and EGF-treated samples were further explored in this study. In the first fingerprint (Fig. 1A), two differentially amplified products were detected, which had been cloned and sequenced in the course of our previous work. The sizes of these two products are indicated (291 and 317 nt). This fingerprint was used to demonstrate that we could identify differentially regulated genes in an array without isolating, cloning and sequencing the RAP-PCR products. This fingerprint and the second fingerprint in Figure 1, which displayed no differential regulation in response to the treatments, were also used to demonstrate that fainter differentially regulated products not visible on the fingerprint gel could, nevertheless, be observed by the array approach.

The fingerprints in Figure 1 fulfill important criteria of reproducibility. To be suitable for either gel- or array-based analysis, RAP-PCR fingerprints must remain almost identical over an 8-fold dilution of the input RNA. Low quality RAP-PCR fingerprints are usually the consequence of poor control over RNA quality and concentration. Before proceeding with the array hybridization steps, it is wise to verify the high quality of the RAP-PCR step. Because the array method has such high throughput, this extra step is neither costly nor time consuming and can greatly improve efficiency by reducing the number of false positives due to poor fingerprint reproducibility.

RAP-PCR fingerprints chosen from Figure 1 were converted into high specific activity radioactive probes by random primed synthesis using [α-^{32}P]dCTP. For each of the two conditions, EGF-treated and untreated, fingerprints generated from RNA at two different concentrations were converted to probe by random primed synthesis for each of the two different fingerprinting
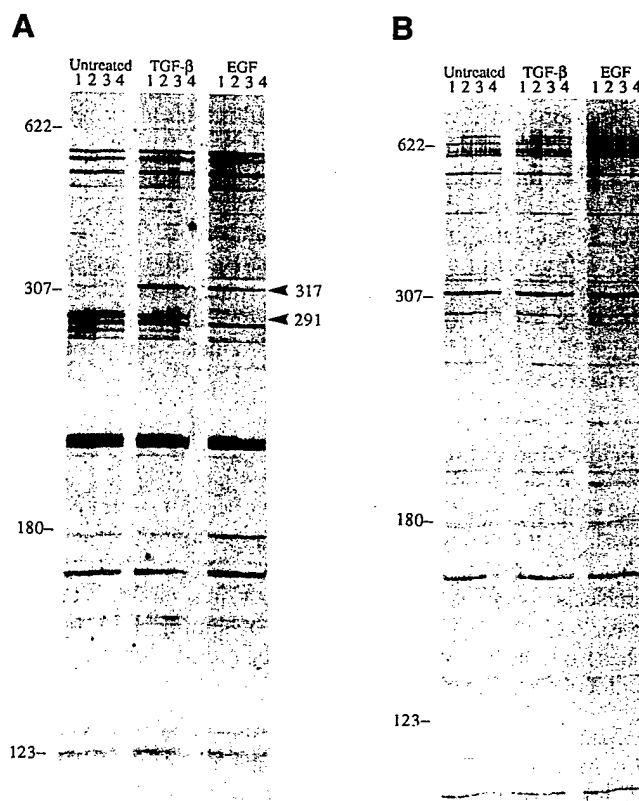

Figure 1. RAP-PCR fingerprints resolved on a gel. Reverse transcription was performed with an oligo(dT) primer on 250, 125, 62.5 and 31.25 ng RNA in lanes 1–4 respectively. RNA was from untreated, TGF-β- and EGF-treated HaCaT cells. RAP-PCR was performed with two sets of primers: (A) GP14/GP16; (B) Nuc1+/OPN24. Molecular weight markers and the sizes of the two differentially amplified RAP–PCR products are indicated.

primer pairs. These radioactively labeled probes were then hybridized to a set of identical arrays each containing 18 432 IMAGE consortium cDNA clones. As controls, total genomic DNA and total poly(A)$^+$ mRNA were also labeled by random priming and used as probes on identical arrays.

Hybridization to arrays

The probes derived from the RAP-PCR fingerprinting reactions described above and the total mRNA and genomic probes were used individually against replicates of a Genome Systems colony array. Hybridization and washing followed standard procedures outlined in Materials and Methods, including the use of genomic DNA as a blocking agent and as a competitor for highly repetitive sequences. Washing at 68°C in 0.1× SSC, 0.1% SDS removed virtually all hybridization to known Alu elements on the membrane, presumably because Alu elements are sufficiently diverged from one another at this wash stringency.

Autoradiograms from the same half of each membrane are shown in Figure 2. Data can also be collected using a phosphorimager, which considerably shortens data collection time and allows quantitation. Other means of labeling, such as fluorescently tagged bases, can be used if suitable arrays and instruments are available. Nylon membranes are typically unsuitable for most commercially
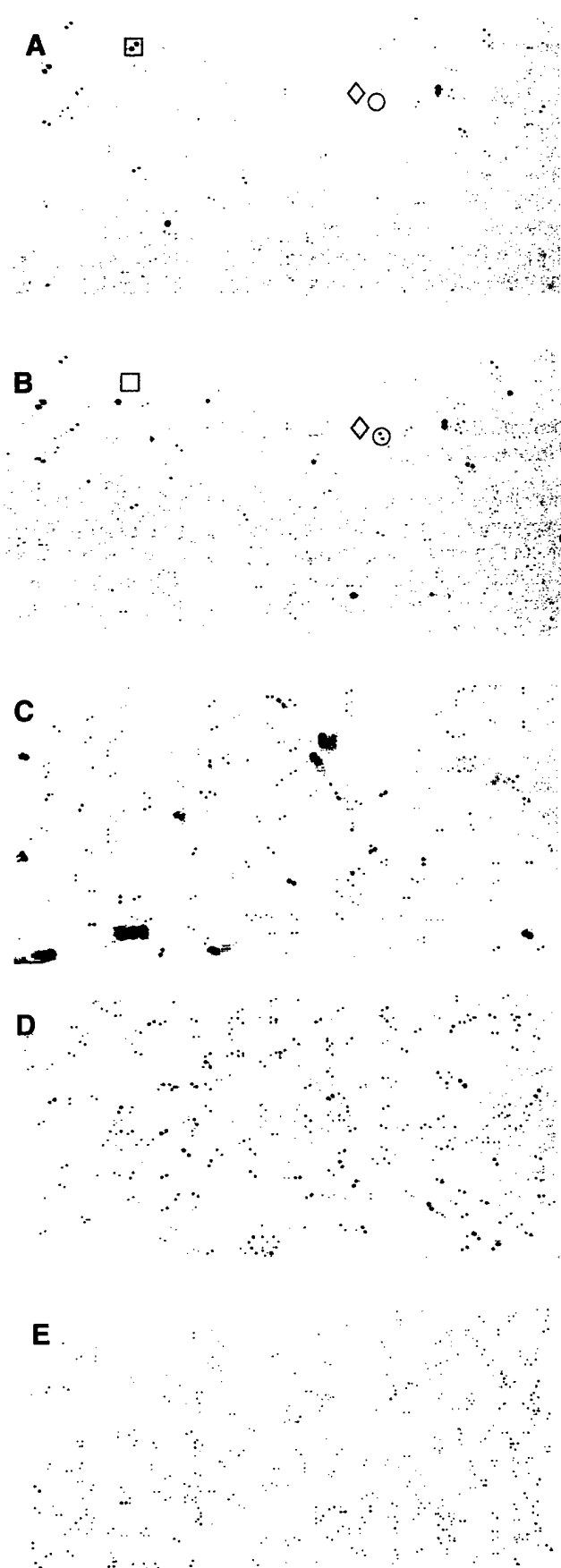
available fluorescent tags due to background fluorescence from the membrane itself.

Overlaps between different probes

The data were analyzed in a number of ways. First, estimates were made of the overlap between the clones hybridized by each probe. In all pairwise comparisons between all of the different types of probes, there was <5% overlap among the 500 clones that hybridized most intensely (compare Fig. 2A, B, D and E). Of the top 500 clones hybridized by the genomic probe (which included nearly all clones known to contain the Alu repeats), <5% overlapped with the top 500 clones hybridized by the fingerprint probes or the total poly(A)$^+$ mRNA probe. This indicated that, except for the case of genomic probe, there was no significant hybridization to dispersed repeats. The overlap among the clones hybridized by the two RAP-PCR fingerprints generated with different primers was <3% and the overlaps of either fingerprint with the poly(A)$^+$ mRNA probe were both <3%. Thus, most of the cDNAs detected using probe from the fingerprints could not be detected using the total mRNA probe. These data indicate that RAP-PCR samples a population of mRNAs largely independently of message abundance. This makes sense because the low abundance class of messages has much higher complexity than the abundant class, making it more likely that the arbitrary primers will find good matches. Unlike differential display, RAP-PCR demands two such arbitrary priming events, possibly biasing RAP-PCR toward the complex class. Overall, these data suggest that the majority of the mRNA population in a cell (<20 000 mRNAs) may be found in as few as 10 RAP-PCR fingerprints.

Further aspects of the data address reproducibility concerns. Using gel electrophoresis, there were no differences among the ~100 bands visible in any of the fingerprints from a single treatment condition performed at different RNA concentrations (Fig. 1). Similarly, >99% of the top 1000 clones hybridized by the probes derived from the fingerprint in Figure 1A were visible at both input RNA concentrations. Furthermore, >98% of the products were the same between the two treatment conditions (i.e. plus and minus EGF) at a single RNA concentration. This indicated almost perfect reproducibility among the top 1000 PCR products in the RAP-PCR amplification.

Figure 2. Hybridization to arrays. All images presented are autoradiograms of the bottom half of duplicates of the same Genome Systems filter probed with radiolabeled DNA. (**A**) and (**B**) Two RAP-PCR reactions using the same primers; (A) untreated; (B) EGF-treated. Three double-spotted clones that show differential hybridization signals are marked on each array. The GenBank accession nos of the clone and the corresponding genes are: square, H10045 and H10098, *vav-3*, AF067817 (13); circle, H28735, gene unknown, similar to heparan sulfate 3-*O*-sulfotransferase-1, AF019386 (17); diamond, R48633, gene unknown. A >10-fold down-regulation for *vav-3*, a >10-fold up-regulation for H28735 and an ~3-fold up-regulation of R48633 were independently confirmed by RT-PCR. (**C**) RAP-PCR using the same RNA as in (A) but with a different pair of primers yields an entirely different pattern. (**D**) cDNA, generated by reverse transcription of 1 µg poly(A)$^+$-selected mRNA. (**E**) Human genomic DNA labeled using random priming.
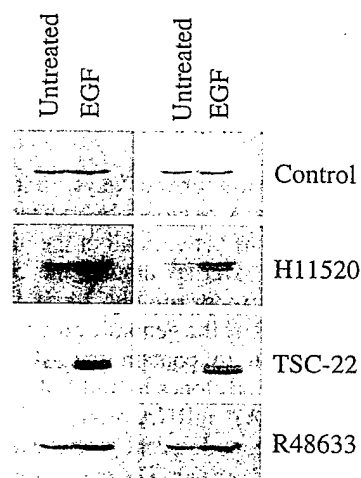
Figure 3. Confirmation of differential regulation by low stringency RT–PCR. Reverse transcription was performed at two RNA concentrations (500 ng, left column; 250 ng, right column). The reaction was diluted 4-fold in water and one fourth was used for low stringency RT–PCR at different cycle numbers. Shown are the control bands, the bands for GenBank accession no. H11520 (both at 22 cycles), the bands for TSC-22 [H11073 and H11161 (27–29)] and the bands for R48633 (all at 19 cycles). H11520 and TSC-22 are ~8–10-fold up-regulated by EGF. R48633 is ~3-fold up-regulated.

The detection of differentially regulated genes using RAP-PCR-derived probes against cDNA arrays

These experiments were designed to detect genes differentially regulated by EGF and TGF-β treatment in confluent keratinocytes. The fingerprint in Figure 1A reveals two boldly differentially regulated genes, the sequences of which were determined during the course of previous work (data not shown). The choice of which Genome Systems arrays to use was based on the presence of these clones. Figure 2 shows the results of hybridization of probes from these fingerprints to the arrays. Arrayed clones corresponding to the 291 nt (*vav-3*, square) and 317 nt (similar to N-HSST, circle) sequenced RAP-PCR fragments are indicated (compare Fig. 2A and B).

Also indicated on this array is a differentially regulated gene that could not be visualized on the original fingerprint gel. This result indicates that differential gene regulation can be detected by the combined fingerprinting and array approach even when the event cannot be detected using the standard gel electrophoresis approach. Verification of differential expression was performed by RT–PCR and will be described in the next section.

A total of 30 differentially hybridizing cDNA clones were detected among ~2000 hybridizing colonies using probes derived from both sets of arbitrary primers (Fig. 1) at a threshold of ~3-fold differential hybridization. Twenty two of these differentially hybridizing clones displayed differential hybridization at both RNA concentrations. These 22 were carried further to the RT–PCR confirmation step, described in the next section.

The eight false-positive clones that appeared to be regulated at only one concentration were of interest in exploring sources of error in the system. Of these eight, five potentially miscalled cases showed differential hybridization at one concentration but were present and not regulated on the membranes for the other

concentration. The most likely source of this type of error is in the membranes. Although each clone is spotted twice, it is possible that occasionally one membrane received substantially more (or less) DNA in both spots than the other three membranes for these clones. However, this potential error was easily detected and is rare, occurring only five times in >2000 clones. The other three potentially miscalled cases hybridized under only one treatment condition and at only one RNA concentration used for RAP-PCR. These may be real differentially expressed genes, but might be false positives from irreproducible PCR products. However, there is an extraordinarily low number of these irreproducible products in the experiments we present here and they are easily identified by comparing the results of two probes derived from PCR of different starting concentrations of RNA.

Confirmation of differential expression using low stringency RT–PCR

Only those hybridization events that indicated differential expression at both input RNA concentrations were carried further. For confirmation of differential expression, we used RT–PCR with specific probes rather than northern blots because we expected that many of the mRNAs would be rare and northern blots are much less sensitive than RT–PCR. One of the advantages of using the arrays from the IMAGE consortium is that >70% of the clones have single pass sequence reads from the 5′- or 3′-end or both deposited in the GenBank database. Thus, it is usually not necessary to sequence a clone in order to derive primers for specific PCR. In cases where there is no sequence available, the clones can be ordered from Research Genetics and sequenced. We have used this strategy in the past, but in this report we confine ourselves to clones for which some sequence is available in the database. Five of the 22 ESTs representing differentially regulated genes on the array had not been sequenced and two of the remaining 17 ESTs were from the same gene. This left 15 unique sequenced genes. In all cases we attempted to align sequences from differentially regulated genes with other sequences in the database in order to derive a higher quality sequence from multiple reads and longer sequence from overlapping clones. The UniGene database clusters human and mouse ESTs that appear to be from the same gene (10). This database greatly aids in the process of assembling a composite sequence from different clones of the same mRNA (http:// www.ncbi.nlm.nih.gov/UniGene/ index.html). These composite sequences were then used to choose primers for RT–PCR.

For each gene, two specific primers were used in RT–PCR under low stringency conditions similar to those used to generate RAP-PCR fingerprints. In addition to the product of interest, a pattern of arbitrary products is generated which is largely invariant and behaves as an internal control for RNA quality and quantity and for reverse transcription efficiency (11). The number of PCR cycles was adjusted to between 14 and 25 cycles, according to the abundance of the product, in order to preserve the differences in starting template mRNA abundances. This is necessary because rehybridization of abundant products during the PCR inhibits their amplification and the difference in product abundances diminishes as the number of PCR cycles increases, in what we have called the 'Cot effect' (12).

Low stringency RT–PCR experiments confirmed the differential expression of the two transcripts that were identified in the RAP-PCR fingerprints of Figure 1A and showed differential

hybridization to the cDNA array (Fig. 2A versus B). These genes had previously been isolated from the gel in Figure 1 and sequenced. One of these corresponds to a new family member of the *vav* proto-oncogene family (13–16) and the other has homology to heparan sulfate 3-*O*-sulfotransferase-1 (17). These have been shown to be regulated under a variety of experimental conditions (manuscript in preparation). The other 13 candidates were also tested and 11 were confirmed. Examples are presented in Figure 3. A list of these genes is given in Table 1. Of the two that were not confirmed, one proved unamplifiable, perhaps because of the low quality sequence used to make the primers or because hybridization to the array was by a differentially regulated closely similar family member. The other gene gave a product but appeared not be differentially regulated by RT–PCR. In addition to the possibility of a family member being regulated, this result could also be due to differential processing of the mRNA rather than differential promotor activity. There is already a precedent for this: differential processing appears to be the reason that *vav-3* yields differential hybridization and differential processing is only observed if the correct primers are chosen (manuscript in preparation).

Table 1. Genes regulated by >2-fold after EGF treatment of confluent HaCaT keratinocytes[a]

Accession number	Gene name
Up-regulated	
H11520 (3′)	Unknown
H11161 (5′)/H11073 (3′)	TSC-22 [U35048]
R48633 (5′)	Unknown
H28735 (3′)	Similar to heparan sulfate 3-*O*-sulfotransferase-1 precursor [AF019386]
H25513 (5′)/H25514 (3′)	Fibronectin receptor α subunit [M13918]
H12999 (5′)/H05639 (3′)	Similar to focal adhesion kinase (FAK2) [L49207]
H15184 (5′)/H15124 (3′)	*ray* gene [X79781]
H25195 (5′)/H24377 (3′)	X-box binding protein-1 (XBP-1) [M31627]
H23972 (3′)	Unknown
H27350 (5′)	CPE receptor (hCPE-R) [AB000712]
R75916 (5′)	Similar to semaphorin C [X85992]
Down-regulated	
R73021 (5′)/R73022 (3′)	Epithelium-restricted Ets protein ESX [U66894]
H10098 (5′)/H10045 (3′)	*vav-3* [AF067817]

[a]Differential expression was confirmed by low stringency RT–PCR. The left column gives the accession numbers of the EST clones (5′ or 3′ or both when available). The right column gives the corresponding gene or the closest homolog. In cases of very low homlogies the gene is considered unknown.

Detecting rare mRNAs

How effective are RAP-PCR probes at detecting rarer mRNAs? Each fingerprint hybridizes to a set of clones almost entirely different from the set hybridized by a probe derived from poly(A)[+]-selected mRNA (Fig. 2). In addition, numerous other primer pairs, membranes and sources of RNA consistently show a <5% overlap between clones hybridized by any two fingerprints or between a fingerprint and a total poly(A)-selected cDNA probe (data not shown). We also attempted to use a northern blot of poly(A)-selected RNA to detect the *vav-3* mRNA (Fig. 1A), which is a new member of the *vav* oncogene family. Despite our ability to detect serially diluted vector down to the equivalent of a few copies per cell, we were unable to detect *vav-3* mRNA, whereas RT–PCR confirmed expression. A glyceraldehyde 3-phosphate dehydrogenase control indicated that the northern blot was performing correctly (data not shown). *vav-3*, therefore, appears to be a low abundance message that is represented in a RAP-PCR fingerprint as a prominent band.

The frequency of homologs of cDNAs detected by the RAP-PCR probes in the EST database was determined (>98% identity). This was compared with the frequency of homologs for a random set of other cDNAs on the same membrane. If the RAP-PCR fingerprints were heavily biased towards common mRNAs, then many would occur often in the EST database because it is partly derived from cDNA libraries that are un-normalized or incompletely normalized. However, the cDNAs detected by RAP-PCR had frequencies in the EST database comparable with the frequencies for randomly selected cDNAs, including cases where the clone was unique in the database. This implies that sampling by RAP-PCR is at least as good as random sampling of the partly normalized libraries used to construct the array, and certainly very different from that obtained for an un-normalized probe such as total mRNA.

Comparison with other sampling methods

In principle, there are several ways to generate a reduced complexity cDNA probe. One of the most successful ways to reduce probe complexity while accentuating the differences between two samples is to perform subtraction (see for example 18), which can have a sensitivity of 1/200 000 (19). It is an obvious but important extension of this manuscript that it would be worthwhile to screen mixtures resulting from subtraction using arrays of ESTs or total cDNAs, when they become available.

Subtraction can be applied to RAP-PCR by simply quenching a labeled fingerprint with an unlabeled fingerprint and we have preliminary evidence that this works (data not presented). A limitation of subtraction is that it can eliminate differences that fall short of presence versus total absence of a mRNA. Furthermore, while subtraction is useful in a binary question, it is of limited utility in cases where a large number of conditions are to be compared combinatorially.

There are two fundamentally different types of complexity reduction: those that maintain the relative stoichiometry among the mRNAs they sample, and those which do not. In the former category are strategies such as selecting a narrow size class of mRNAs or cDNAs (20), where rare mRNAs would still be rare. Other methods that maintain approximate stoichiometry include those that employ 3′-anchored cDNA restriction fragments (see for example 21–23). In a RAP-PCR both the abundance and the degree of match with the primers contribute to the prevalence of any particular product. Thus, rare mRNAs that happen to have excellent matches with the primers and are efficiently amplified are found among the more abundant RAP-PCR products. In this respect, a RAP-PCR probe is non-stoichiometric (24). This is a

very useful feature of RAP-PCR because it allows the sampling of mRNAs that are difficult to sample using other methods.

Changing the number of products sampled by RAP-PCR

Detection is ultimately limited by background hybridization and incomplete blockage of repeats. At present, ~1000 cDNAs on a colony array of 18 432 clones can be reliably scored by each RAP-PCR probe and the limitation seems to be the number of sufficiently abundant products in the PCR reaction rather than background.

The effect of RAP-PCR reaction parameters on the distribution and number of products that can be observed on arrays has not been fully explored, including the optimal complexity of the probe. To increase the complexity we used *Taq* polymerase Stoffel fragment, which is more promiscuous than Ampli*Taq*. The primers used were 10 or 11 bases in length and are not degenerate (they have a single base at each position). Longer primers used at the same temperature might give a more complex product, as would primers with some degeneracy. We have recently used an oligo(dT) primer anchored at the 3′-end as one of the two primers (manuscript in preparation). Anchoring at the 3′-end of messages (25) should result in more hybridization in arrays that are 3′-biased. However, the greater the complexity of the probe, the more closely it will resemble a total mRNA probe, which loses the advantage of non-stoichiometric sampling. Using arrays will teach us a number of things about the RAP-PCR mixtures that would be much less evident from a gel. For example: How complex are the fingerprints? What is the distribution of products among those easily seen on a gel and those that are too rare to be seen on a gel? What are the various effects of primer length, degeneracy and anchoring in the reverse transcription and PCR reactions? What are the effects of various different polymerases at each of these steps? Some of the answers to these questions will undoubtedly improve the throughput of the method for arrays.

Colony arrays used in these experiments represent the worst case scenario, in which plasmid DNA is mixed with a large mass excess of a bacterial genome having 5 Mb of complexity. If probes become so complex that background becomes the limiting factor, more sophisticated arrays may become essential. PCR product arrays or oligonucleotide arrays may yield higher scorability with more complex probes.

Comparison with gel-based characterization of fingerprints

Cloning genes from RAP-PCR fingerprints resolved on gels may still have some advantages over using fingerprints for probes on arrays in certain circumstances. For example, a new gene could be found that is not already on a membrane. However, this advantage diminishes every day as more cDNAs are characterized. Another advantage is that close family members can yield different PCR products in fingerprints, whereas on arrays a close family member may hybridize to a clone and lead to a misinterpreted result. This possibility is somewhat diminished if 3′ ESTs are used on membranes, because 3′-ends of even quite closely related genes may be sufficiently divergent to avoid that problem. Furthermore, fingerprints resolved on gels may detect new splicing variants, which is less likely using clone arrays. Also, when using gel-based fingerprints a single primer pair can be used to survey 100 mRNAs in a very large number of RNA

samples. The same number of fingerprints applied to arrays would be expensive, though this is balanced by the fact that 1000 mRNAs would be surveyed. Overall, for most applications, the ability to screen many thousands of genes with a single fingerprint on a series of arrays will outweigh the advantages of gel-based assays. In addition, any bias toward abundant transcripts that exists in RAP-PCR is partially mitigated by the array approach, because even products that are never visible in the gels can still serve as effective probes.

The rate of throughput using fingerprints as probes for arrays compares favorably with that obtainable using gel fingerprints. A single sequencing-style gel loaded with RAP-PCR fingerprints from 25 different primer pairs usually displays ~1000 products. This is similar to the number of mRNAs we surveyed on a single membrane containing about one third of the unique sequences in the IMAGE consortium clones. A fingerprinting experiment using 25 primer pairs could supply probes for cDNA arrays conceivably covering >20 000 genes.

Hybridization of fingerprints to arrays has the huge advantage that there is generally no need to isolate, clone and sequence the genes detected. In principle, all known human mRNAs will fit on three membranes (~50 000 genes). At present, each fingerprint has sufficient complexity to hybridize to >2000 of the 50 000 known genes. There is also the issue of diminishing returns. In a fingerprint, one cannot know if a differentially amplified product has been sampled previously without performing considerable further work. In contrast, with an array one always knows what genes have been sampled previously. In principle, one can even select primers that enrich for genes not yet sampled (26).

In summary, a method is presented in which the intrinsic reduced complexity and non-stoichiometric amplification resulting from arbitrarily primed PCR fingerprinting is used to advantage to construct probes for cDNA arrays. Simple methods that allow inexpensive arrays to generate useful information are likely to allow many molecular biology laboratories to participate in the revolution in understanding gene regulation that arrays can achieve. We hope that a public resource will soon develop in which the transcriptional effect of a growing list of conditions is attached to every gene. Ultimately, such information will link to the promotors of these genes and to the signal transduction cascades responsible for their regulation.

ACKNOWLEDGEMENTS

REFERENCES

1 Marshall,A. and Hodgson,J. (1998) *Nature Biotechnol.,* **16**, 27–31.
2 Ramsay,G. (1998) *Nature Biotechnol.,* **16**, 40–44.
3 Pietu,G., Alibert,O., Guichard,V., Lamy,B., Bois,F., Leroy,E., Mariage-Sampson,R., Houlgatte,R., Soularue,P. and Auffray,C. (1996) *Genome Res.,* **6**, 492–503.
4 Boll,W., Fujisawa,J., Niemi,J. and Weissmann,C. (1986) *Gene,* **50**, 41–53.
5 Welsh,J. and McClelland,M. (1990) *Nucleic Acids Res.,* **18**, 7213–7218.
6 Welsh,J., Chada,K., Dalal,S.S., Cheng,R., Ralph,D. and McClelland,M. (1992) *Nucleic Acids Res.,* **20**, 4965–4970.

7 Boukamp,P., Popp,S., Altmeyer,S., Hulsen,A., Fasching,C., Cremer,T. and Fusenig,N.E. (1997) *Genes Chromosomes Cancer*, **19**, 201–214.

8 McClelland,M., Ralph,D., Cheng,R. and Welsh,J. (1994) *Nucleic Acids Res.*, **22**, 4419–4431.

9 Mathieu-Daude,F., Trenkle,T., Welsh,J., Jung,B., Vogt,T. and McClelland,M. (1998) *Methods Enzymol.*, in press.

10 Schuler,G.D. (1997) *J. Mol. Med.*, **75**, 694–698.

11 Mathieu-Daude,F., Welsh,J., Davis,C. and McClelland,M. (1998) *Mol. Biochem. Parasitol.*, **92**, 15–28.

12 Mathieu-Daude,F., Welsh,J., Vogt,T. and McClelland,M. (1996) *Nucleic Acids Res.*, **24**, 2080–2086.

13 Katzav,S., Martin-Zanca,D. and Barbacid,M. (1989) *EMBO J.*, **8**, 2283–2290.

14 Katzav,S. (1995) *Crit. Rev. Oncogen.*, **6**, 87–97.

15 Bustelo,X.R. (1996) *Crit. Rev. Oncogen.*, **7**, 65–88.

16 Romero,F. and Fischer,S. (1996) *Cell Signalling*, **8**, 545–553.

17 Shworak,N.W., Liu,J., Fritze,L.M., Schwartz,J.J., Zhang,L., Logeart,D. and Rosenberg,R.D. (1997) *J. Biol. Chem.*, **272**, 28008–28019.

18 Jin,H., Cheng,X., Diatchenko,L., Siebert,P.D. and Huang,C.C. (1997) *BioTechniques*, **23**, 1084–1086.

19 Rhyner,T.A., Biguet,N.F., Berrard,S., Borbely,A.A. and Mallet,J. (1986) *J. Neurosci. Res.*, **16**, 167–181.

20 Dittmar,G., Schmidt,G., Kopun,M. and Werner,D. (1997) *Cell Biol. Int.*, **21**, 383–391.

21 Bachem,C.W., van der Hoeven,R.S., de Bruijn,S.M., Vreugdenhil,D., Zabeau,M. and Visser,R.G. (1996) *Plant J.*, **9**, 745–753.

22 Habu,Y., Fukada-Tanaka,S., Hisatomi,Y. and Iida,S. (1997) *Biochem. Biophys. Res. Commun.*, **234**, 516–521.

23 Money,T., Reader,S., Qu,L.J., Dunford,R.P. and Moore,G. (1996) *Nucleic Acids Res.*, **24**, 2616–2617.

24 Trenkle,T., Mathieu-Daude,F., Welsh,J. and McClelland,M. (1998) *Methods Enzymol.*, in press.

25 Liang,P. and Pardee,A.B. (1992) *Science*, **257**, 967–971.

26 Pesole,G., Liuni,S., Grillo,G., Belichard,P., Trenkle,T., Welsh,J. and McClelland,M. (1998) *BioTechniques*, **25**, 112–117.

27 Jay,P., Ji,J.W., Marsollier,C., Taviaux,S., Berge-Lefranc,J.L. and Berta,P. (1996) *Biochem. Biophys. Res. Commun.*, **222**, 821–826.

28 Dmitrenko,V.V., Garifulin,O.M., Shostak,E.A., Smikodub,A.I. and Kavsan,V.M. (1996) *Tsitol. Genet.*, **30**, 41–47.

29 Ohta,S., Shimekake,Y. and Nagata,K. (1996) *Eur. J. Biochem.*, **242**, 460–466.